A uniform framework for the objective assessment and optimisation of radiotherapy image quality

Andrew J Reilly

Doctor of Philosophy
The University of Edinburgh
2010

To my family, for encouraging and supporting me thus far,
and to my wife Elaine for the adventures yet to come

It’s not what you look at that matters,
it’s what you see.
Henry David Thoreau

Declaration

I declare that this thesis was composed by myself, that the work contained therein is my own and that it has not previously been submitted as part of the assessment for another degree or professional qualification, except where explicitly stated otherwise in the text.

(*Andrew J Reilly*)

Abstract

Image guidance has rapidly become central to current radiotherapy practice. A uniform framework is developed for evaluating image quality across all imaging modalities by modelling the 'universal phantom': breaking any phantom down into its constituent fundamental test objects and applying appropriate analysis techniques to these through the construction of an automated analysis tree. This is implemented practically through the new software package 'IQWorks' and is applicable to both radiotherapy and diagnostic imaging.

For electronic portal imaging (EPI), excellent agreement was observed with two commercial solutions: the QC-3V phantom and PIPS Pro software (Standard Imaging) and EPID QC phantom and epidSoft software (PTW). However, PIPS Pro's noise correction strategy appears unnecessary for all but the highest frequency modulation transfer function (MTF) point and its contrast to noise ratio (CNR) calculation is not as described. Serious flaws identified in epidSoft included erroneous file handling leading to incorrect MTF and signal to noise ratio (SNR) results, and a sensitivity to phantom alignment resulting in overestimation of MTF points by up to 150% for alignment errors of only ±1 pixel.

The 'QEPI1' is introduced as a new EPI performance phantom. Being a simple lead square with a central square hole it is inexpensive and straightforward to manufacture yet enables calculation of a wide range of performance metrics at multiple locations across the field of view. Measured MTF curves agree with those of traditional bar pattern phantoms to within the limits of experimental uncertainty. An intercomparison of the Varian aS1000 and aS500-II detectors demonstrated an improvement in MTF for the aS1000 of 50–100% over the clinically relevant range 0.4–1 cycles/mm, yet with a corresponding reduction in CNR by a factor of $\sqrt{2}$. Both detectors therefore offer advantages for different clinical applications.

Characterisation of cone-beam CT (CBCT) facilities on two Varian On-Board Imaging (OBI) units revealed that only two out of six clinical modes had been calibrated by default, leading to errors of the order of 400 HU for some modes

and materials – well outside the ±40 HU tolerance. Following calibration, all curves agreed sufficiently for dose calculation accuracy within 2%. CNR and MTF experiments demonstrated that a boost in MTF f_{50} of 20–30% is achievable by using a 512^2 rather than a 384^2 matrix, but with a reduction in CNR of the order of 30%.

The MTF f_{50} of the single-pulse half-resolution radiographic mode of the Varian PaxScan 4030CB detector was measured in the plane of the detector as 1.0±0.1 cycles/mm using both a traditional tungsten edge and the new QEPI1 phantom. For digitally reconstructed radiographs (DRRs), a reduction in CT slice thickness resulted in an expected improvement in MTF in the patient scanning direction but a deterioration in the orthogonal direction, with the optimum slice thickness being 1–2 mm. Two general purposes display devices were calibrated against the DICOM Greyscale Standard Display Function (GSDF) to within the ±20% limit for Class 2 review devices.

By providing an approach to image quality evaluation that is uniform across all radiotherapy imaging modalities this work enables consistent end-to-end optimisation of this fundamental part of the radiotherapy process, thereby supporting enhanced use of image-guidance at all relevant stages of radiotherapy and better supporting the clinical decisions based on it.

Table of Contents

List of Figures

List of Tables

Abbreviations

ADC Analogue to Digital Converter

aSi amorphous Silicon

BEV Beam's Eye View

BKE Background Known Exactly

CAX Radiation Beam Central AXis

CBCT Cone-Beam CT

CEP Centre for Evidence-based Purchasing

CNR Contrast-to-Noise Ratio

CoE Centre of Extremes

CoM Centre of Mass

CoV Coefficient of Variation

CRT Conformal Radiotherapy

CSV Comma Separated Variable

CT X-Ray Computed Tomography

CTDI CT Dose Index

CTV Clinical Target Volume

DDL Digital Driving Level

DICOM Digital Imaging and Communications in Medicine

DLP Dose Length Product

DQE Detective Quantum Efficiency

DR Digital Radiology

DRR Digitally Reconstructed Radiograph

dSNR Detail or Difference Signal-to-Noise Ratio

DVHs Dose Volume Histograms

ECC Edinburgh Cancer Centre

EPI Electronic Portal Imaging

EPID Electronic Portal Imaging Device

FFT Fast Fourier Transform

FWHM Full-Width Half-Maximum

GSDF Grayscale Standard Display Function

GTV Gross Tumour Volume

HU Hounsfield Unit

ICRU International Commission on Radiation Units and Measurements

IGRT Image-Guided Radiotherapy

IMRT Intensity Modulated Radiotherapy

IRF Impulse Response Function

ISL Inverse Square Law

ITK Insight Toolkit

JND Just-Noticable Difference

LDPE Low Density Polyethylene

LSF Line Spread Function

LSI Linear and Shift-Invariant

LUT Look-Up Table

MLC Multi-Leaf Collimator

MRI Magnetic Resonance Imaging

MTF Modulation Transfer Function

MU Monitor Unit

NEA Noise Equivalent Aperture

NEQ Noise Equivalent Quanta

NIH National Institutes of Health

NNPS Normalised Noise Power Spectrum

NPS Noise Power Spectrum

NPWMF Non Pre-Whitening Matched Filter

NTCP Normal Tissue Complication Probability

OAR Organ at Risk Volume

OCC Oxford Cancer Centre

OOD Object-Orientated Design

OTF Optical Transfer Function

PACS Picture Archive and Communication System

PET Positron Emission Tomography

POM Polyoxymethylene

PPM Planned Preventative Maintenance

PRV Planning at Risk Volume

PSF Point Spread Function

PTFE Polytetrafluoroethylene

PTV Planning Target Volume

PWMF Pre-Whitening Matched Filter

QA Quality Assurance

QC Quality Control

ROC Receiver-Operator Characteristic

ROI Region of Interest

RV Record and Verify

SAD Source-Axis Distance

SD Standard Deviation

SDT Signal Detection Theory

SKE Signal Known Exactly

SNR Signal-to-Noise Ratio

SSD Source to surface distance

TCP Tumour Control Probability

TFT Thin Film Transistor

US Ultrasound

XML Extensible Markup Language

Acknowledgements

Many people offered support, valuable advice and access to resources throughout the course of this extended project and I am indebted to them all.

First of all, I'd like to thank my supervisors David Thwaites and Tony Redpath for their support, encouragement – and patience – as this project progressed, and for engaging in helpful discussions right up until the end.

I'd also particularly like to thank my colleagues in Edinburgh and Oxford who were involved in the development and testing of IQWorks and in performing many of the routine quality assurance measurements that I am able to report on in this thesis. Special thanks are due to Dave Colburn and Bryan Doig for preparing early prototypes of the QEPI1 portal imaging phantom in Edinburgh, and to Terry Wolf and Steve Jones for manufacturing the production versions currently in use in Oxford. Thank you also to Steve Mutch, Mary Cocker and Nick Weir for guidance and help in performing CT dosimetry measurements.

I'd also like to express gratitude to Sarah Muscat and the Beatson Oncology Centre for the loan of their RMI-467 phantom to perform electron density measurements in Edinburgh, and to Phil Neal (Imaging Equipment Ltd), Dawn Broadhead (QADOS) and Colour Confidence Ltd for the loan of various phantoms and measurement instruments over the years in support of this work.

Whilst undertaking this project I've kept in touch with developments by attending scientific meetings at home and abroad. I am very grateful to the Institute of Physics and Engineering in Medicine (IPEM), the Institute of Physics (IOP) and both the Edinburgh and Oxford physics departments for supporting me in these. In particular I'd like to thank IPEM for the generous Travel Award which enabled me to spend a month in the USA visiting departments working at the forefront of introducing diagnostic imaging methodology into radiotherapy physics. Thanks are due also to those departments for accommodating me.

IQWorks is a collaborative project and I am grateful to the support I have received form the wider radiotherapy physics and diagnostic imaging communities. In particular, I'd like to thank Ed McDonagh, Elly Castellano, David Platten and Nick Marshall for their enthusiasm, for being instrumental in taking

this project forward, for numerous invaluable discussions and for supplying data and giving up their precious time to help validate the code. I'm also very grateful to the CT User Group for continuing to sponsor the IQWorks website.

Undertaking a PhD part time is a challenging venture. I'm indebted to the Edinburgh and Oxford departments for supporting me in my studies through giving me time and access to resources. In particular, I'd like to thank my managers Tony Redpath (Edinburgh), Elizabeth Macaulay (Oxford) and Chris Gibson (Oxford) for their help and support. I'm also very grateful to Tony Redpath and Allan Price, Professor of Radiation Oncology at the University of Edinburgh, for contributing financially to my university fees: your help was very much appreciated.

Working towards a PhD impinges not just upon the workplace but upon all aspects of life. A very special thank you is due to my parents, Vicky and Jim, and sister Anne for their support and encouragement over the years, and to all my friends who have travelled on this journey with me. Carolyn, thank you especially for showing me it was possible!

Finally, I'd like to thank my wife Elaine for her unflinching support during the write-up and final phases of this work, and for helping me keep focused on the light at the end of the tunnel.

Chapter 1.

Introduction

1.1. Overview

Over 285,000 new cancers are diagnosed in the United Kingdom each year, with 1 in 3 people developing cancer over the course of their lifetimes. Currently, 1 in 4 of all deaths can be attributed to cancer.

Treatment of cancer typically involves excision of a tumour by surgery, attempting to arrest progression of the disease using pharmaceuticals ('chemotherapy') or attacking the tumour using ionising radiation ('radiotherapy'). Most treatments involve some combination of these, with it being optimal to employ radiotherapy in over half of all cases. Today, around 50% of all patients diagnosed with a new cancer will survive for 5 years, with radiotherapy contributing to the cure for 40% of these. Radiotherapy is also widely utilised as an efficient, non-invasive means of palliation, relieving pain through symptomatic control[23, 61, 66, 91, 259, 261].

Radiotherapy is therefore an established and proven weapon in the fight against cancer. Almost all curative – and many palliative – radiotherapy treatments are underpinned by some form of medical imaging, with the extent to which imaging is relied upon depending upon the complexity of the treatment. As treatment techniques become increasingly sophisticated the imaging supporting the radiotherapy process also increases significantly, along with the complexity of the studies required and the range of modalities employed[23, 155, 162, 188, 261, 287, 288, 316].

This chapter discusses the evolving role of imaging in radiotherapy, the ongoing drive towards treatment optimisation and the requirement for a common framework for the objective assessment of image quality in order to facilitate this. The thesis that it is possible to apply a uniform methodology across all imaging modalities is introduced and the goals of this work are described.

1.2. Radiotherapy Technology

Radiotherapy involves delivering a high dose of ionising radiation to an identified tumour volume, with the goal of destroying the tumour whilst minimising damage to nearby healthy tissues[145, 154]. Dose can be delivered by directly placing a radioactive source inside the volume to be treated (known as *brachytherapy*, from the Greek for 'near' therapy). Alternatively, beams of radiation can be fired in from outside the body, a technique referred to as *external beam radiotherapy* or *teletherapy* (from the Greek for therapy 'at a distance'). External beam radiotherapy most commonly utilises X-ray photons or electrons, individually or in combination. Protons and carbon ions can also be employed, but the equipment required to deliver these is considerably more expensive. For this reason, although these technologies offer distinct advantages for certain tumour sites[35, 325], they are currently limited to only the largest clinical centres and research sites [165, 166].

Although worldwide there is active research and development in all modes of radiotherapy, and imaging plays a key role in each of these, the work presented in this thesis concentrates almost exclusively on external beam radiotherapy with X-ray photons. Being by far the most commonly encountered radiotherapy modality it is the one for which the techniques described have been primarily developed and tested. However, all the methods and conclusions presented are directly applicable to other radiotherapy technologies, as well as in the sphere of diagnostic imaging[296].

External beam X-ray therapy is delivered using electron linear accelerators ('linacs'), a schematic diagram of which is shown in Figure 1.1. Most of the work described was undertaken using Varian Clinac 600 and 2100 series linacs, so the description which follows is biased towards these, but all medical linacs follow the same basic operating principals. Using microwaves, electrons are accelerated to megavoltage energies then are collided with a metal target. Although the acceleration path is 'linear', in that the energy gain is achieved by passing the beam through microwave fields in a successive series of waveguide cavities, to align this along a single straight line may result in an impractically long physical system. Bending magnets may be therefore be employed to separate the length along which acceleration takes place from the shorter path to the target, as illustrated in the diagram. The exact configuration depends upon the design of and technologies employed in the linac.

When the electrons impinge upon the target they rapidly decelerate, releasing energy as heat and *bremsstrahlung* (or 'braking') X-rays, with a peak energy

the same as that of the incident electron beam. The result is an X-ray beam emanating from approximately a point source in the target, with the intensity of X-ray photons greatest in the direction parallel to the incident electron beam and reducing radially away from this. To achieve a uniform dose profile under reference conditions (10 cm deep for a 10×10 cm^2 field in a full-scatter water phantom) a metal 'flattening filter' is added to the beamline which preferentially removes photons from the centre of the beam where the fluence is more intense.

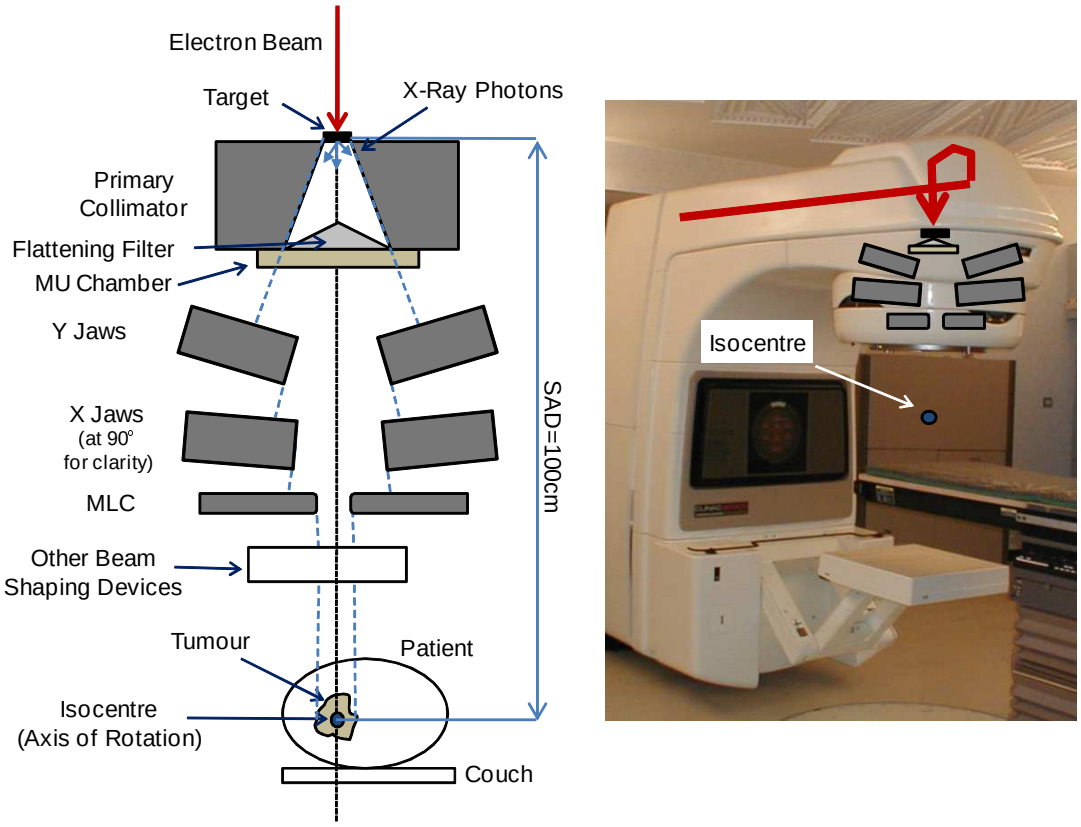


Figure 1.1. – Left: Schematic diagram of a Varian linac, illustrating progressive collimation of the X-ray beam. Right: Photo of 2100CD clinac with key head features superimposed.

Collimation of the beam is provided first by the tungsten housing surrounding the target (the primary collimator), then by two pairs of secondary 'jaws' moving perpendicular to each other (defined 'X' and Y'), resulting in a rectangular shaped radiation field. In *conformal radiotherapy* (CRT), which is the current baseline standard for radical (curative) treatments, additional beam-shaping devices are used to fine-tune the beam portal to the shape of the target. Traditionally, these were shaped lead blocks, mounted on trays beneath the head of the linac, but multi-leaf collimators (MLCs) are increasingly common in contemporary machines. These advanced collimation devices consist of

thin tungsten leaves which can individually and independently be inserted by different amounts into the radiation field, achieving a similar effect to lead blocks but having the advantage that the field-shaping can be quickly adjusted, or turned on or off, using a computer control system. In Varian linacs, the MLC is mounted as a tertiary collimator below the X jaws, with the 'Millennium' MLCs used in this study consisting of 60 pairs of leaves that provide collimation in steps of 5 mm over the central 20 cm of the beam and 10 mm in the periphery beyond this[349]. Other manufacturers employ their own proprietary MLC arrangement, either as a tertiary collimator or by substituting one of the secondary jaw pairs[43, 87, 88].

All collimation devices can rotate around the beam central axis (CAX) so that optimum conformation to irregular tumour shapes can be achieved. In addition, the linac and patient couch also rotate about a single point in space, centred on the beam CAX. By placing the target at this point of mutual rotation (the 'isocentre') it is possible to treat with beams from multiple angles, always hitting the target but entering and exiting through different over- and underlying volumes of healthy tissue each time, thus spreading out unwanted dose. International convention is that the distance from the X-ray source to the axis of rotation (SAD) is 100 cm[41].

Radiation dose delivered in radiotherapy is specified as 'absorbed dose' in gray (Gy), which is the joules of energy deposited per kilogram of material. However, even for straightforward treatments, the relationship between the quantity of radiation output by the linac, and the dose delivered at depth in a medium is relatively complex, depending upon the beam energy, depth of measurement, source to surface distance (SSD), collimation and other factors. An ionisation chamber positioned in the head of the linac downstream of the flattening filter provides an internal reference for measuring linac output. Completely encompassing the radiation beam, this monitor chamber is sensitive to changes in beam characteristics and linac performance. Output is measured in terms of *monitor units* (MU), usually an integral number, with chambers calibrated so that a given number of MU corresponds to a particular absorbed dose under reference conditions. The relationship between MU and absorbed dose at depth in a uniform medium for any geometrical configuration can then be determined using look-up tables of correction factors. Most radiotherapy centres, including those in Edinburgh and Oxford, calibrate the monitor chambers so that delivering 100 MU using a 10×10 cm^2 square field to a water phantom at SSD 100 cm, corresponds to an absorbed dose along the beam CAX

of 1 Gy at the depth of dose maximum[365].

The X-ray energy spectrum of the photon beam depends upon a number of factors, including the range of energies in the original electron beam and the physical composition and geometry of the target and flattening filter. Beams are generally specified as an energy in megaelectronvolts (MeV), although this is usually written as MV to indicate there is actually a range of photon energies present. The energy value roughly corresponds to the maximum energy in the photon spectrum, but there can be confusion because different linac beamlines may result in the same peak energy but significantly different spectra. To help standardise practice, depth-dose curves of reference beam energies are tabulated in BJR Supplements 17 and 25[27, 28], and the energy name assigned to an actual beam is usually that of the closest matching curve in these reports. To avoid all ambiguity in the clinic, beam energy is characterised by physical measurement of the *quality index*, which is the ratio between the absorbed dose deposited at two depths in a water phantom, with the point of measurement always being at the SAD[205]. Two beams from similar beamlines and with the same quality index can be taken as having the same depth-dose characteristics.

X-ray based radiotherapy is successful because electronic equilibrium conditions – when the flux of secondary electrons scattered into an elemental volume is equal to that being scattered out – are not reached until a beam has penetrated to depth within the patient. In the 'build-up' region before this happens the dose increases with depth as the kinetic energy lost by scattered electrons is gradually balanced out. Once equilibrium is established, the dose then falls-off with depth according to the Inverse Square Law (ISL) and as a result of absorption interactions. A large dose can be therefore be delivered to deep tumours whilst sparing superficial tissues. The actual depth of dose-maximum (d_{max}) increases with beam energy.

State-of-the-art linacs are entirely computer controlled and are capable of performing all motions of the linac gantry, couch and collimation devices either automatically or semi-automatically. The 'Record and Verify' (RV) system checks all initial parameters are set according to a pre-determined plan before allowing the treatment to commence, prohibiting dose delivery if there are any discrepancies, although many such *interlocks* can be overridden if appropriate by an authorised operator. The system subsequently records all overrides and the actual parameters used to deliver the treatment, including the quantity of radiation delivered[180]. Very advanced dynamic treatments, such as intensity-modulated radiotherapy (IMRT), are possible through combinations

of simultaneous motions of the jaws and MLC leaves, rotating the gantry and modulating the radiation dose-rate whilst the beam is being delivered[20, 207].

1.3. Clinical Radiotherapy

Management of a patient's cancer is a complex, multi-faceted process involving numerous professional disciplines and clinical teams. Radiotherapy constitutes only one part of this process[134, 288, 339], as illustrated in Figure 1.2. Once a cancer diagnosis has been confirmed, and radiotherapy has been selected as one of the treatment modalities, responsibility passes to the radiotherapy team for that part of the treatment.

Radiotherapy is generally a sequential process, traditionally divided into treatment preparation ('pre-treatment') and treatment phases. First, a treatment plan is developed to achieve the specified clinical goals. This is prepared by a multi-disciplinary team consisting of clinicians, radiographers, technologists and physicists and will involve imaging studies, optimisation of beam parameters (angles, collimation, energy, relative weighting, etc.), dose calculations and verifying the treatment is practically deliverable. The treatment is then delivered over a number of sessions (or *fractions*), taking advantage of tumour cells recovering more slowly than normal cells after each session and allowing a higher dose to be delivered overall[286].

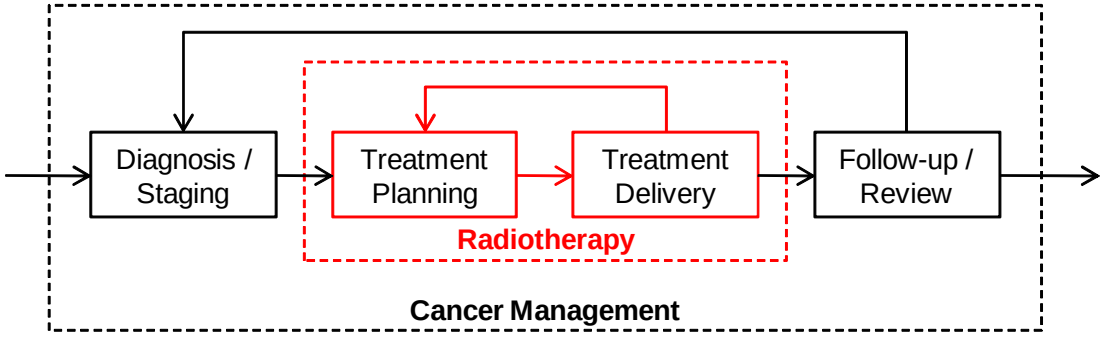


Figure 1.2. – The simplified radiotherapy pathway.

However, with advanced treatment strategies and increasingly integrated computer systems, the distinction between the two phases is becoming less, and it is not uncommon for elements of the treatment plan to be refined as the fractionation course progresses. Because information prepared by the pre-treatment team underpins actions during treatment, good communication between the teams responsible for both phases is essential[285]. Furthermore,

if a problem is encountered during treatment, such as the patient's condition deteriorating so that the original plan is no longer achievable, a major revision of the plan may be required before treatment can progress further.

When planning and prescribing a treatment it is crucial to be able to accurately and unambiguously characterise the three dimensional volume to which dose is to be delivered. Internationally accepted guidelines for this have been published by the International Commission on Radiation Units and Measurements (ICRU), in ICRU Report 50[142] and its supplement Report 62[140].

The *Gross Tumour Volume* (GTV) is the tumour which can be identified by clinical examination, either by visual inspection, palpation, or with the help of radiological images. However, there will always be sub-clinical microscopic spread[127] which cannot be detected, regardless of how good the examination process. Because a tumour can potentially regrow from a single remaining clonogenic cell, it is important to deliver a therapeutic radiation dose to this invisible region too. A margin is therefore added around the GTV to take this into account, with the resulting volume being the *Clinical Target Volume* (CTV). Across the GTV→CTV region the tumour cell density gradually decreases, reaching zero at the outer surface of the CTV. However, geometric uncertainties in the delivery process (described in more detail in sections 1.4 and 1.5 below) mean that the CTV may be in a slightly different place each time the patient is set-up for treatment. Involuntary physiological processes such as respiration may also cause intrafractional motion of the CTV. Further margins are therefore added to take account of geometric uncertainties, resulting in the *Planning Target Volume* (PTV), which is then used to develop the treatment plan. By judicial choice of GTV→CTV and CTV→PTV margins, if a plan is prepared which adequately covers the PTV then, in the absence of gross set-up or delivery errors, one can be confident the whole CTV, and thus all tumour cells, are being treated as intended at every fraction.

When planning a treatment it is important to consider the dose delivered to healthy tissues, so that unwanted side-effects can be minimised. Generally, it is not possible to avoid all normal tissue complications, with short-term problems such as skin erythema and longer-term issues such as dryness of mouth being considered acceptable compromises. However, some organs or functional systems are either more radiosensitive, more susceptible to damage, or less able to self-repair than others. These critical structures or systems require special consideration during the planning process and *Organ at Risk* (OAR) volumes may be defined to facilitate this. In a similar way to the CTV, margins are added

to account for geometric uncertainties, leading to *Planning organ at Risk Volumes* (PRVs) which the treatment plan is then tailored to avoid, ensuring doses within these volumes are within acceptable tolerances.

Contemporary treatments are becoming increasingly sophisticated, with a drive towards improved conformation around the target volume[20, 133, 162, 207, 288], thus maximising the dose to the tumour and reducing dose, and hence toxicity, to healthy tissues. Radical (curative) treatment regimes typically involve daily irradiations over a period of up to eight weeks[286].

1.4. Accuracy of Delivery

Almost by definition, the intention in radiotherapy is to deliver a lethal dose to all cancer cells in a tumour, as characterised by the CTV and practically realised over the course of treatment by the nominal PTV. There have been numerous studies investigating the dose-response of tumour cells and those of healthy tissues[36, 81, 241, 247, 268, 321, 364]. Whilst the radiobiological mechanisms of damage are not fully understood, it is generally agreed that the relationship between the likelihood of damage and that of recovery is a steep sigmoidal curve. The curves describing the tumour control probability (TCP) and the normal tissue complication probability (NTCP) are both sigmoidal in nature, and tend to be very close together in terms of relative dose.

It is therefore crucial that the prescribed treatment dose be delivered as accurately as possible to the whole of the defined PTV, with the dose dropping off rapidly beyond this. There are two separate, but related considerations:

1. The dose distribution must be precisely delivered, as planned, to the identified 3D volume in space about the isocentre (i.e. the PTV, although extra margins again are added to account for limitations in beam-shaping capabilities and penumbra at the edges of fields).

2. The spatial volume to which the dose is delivered must accurately match the original anatomical context on which the treatment was planned (i.e. the CTV must really lie within the volume which is being irradiated, and OARs are really being avoided as intended).

For the dose to the target volume to be as intended the dose must be delivered precisely and accurately to the correct volume within the patient's anatomical frame of reference[136, 154]. Geometrical positioning is therefore inextricably linked with correct dose delivery.

Studies have been performed to identify how accurate and precise dose delivery must be. Empirical evidence indicates that an increase in normal tissue dose of around 7% is sufficient to cause a definitely observable clinical reaction[81], whereas a reduction in dose to the target of around 10% is enough to significantly compromise tumour control[241, 247, 268, 321]. Taking these as absolute limits, it is recommended that the total uncertainty in dose at the ICRU specification point is no more than 3% (expressed as 1 standard deviation (SD) of the mean) and that uncertainty elsewhere in the target is no more than 5% (again to 1 SD)[136, 145, 154, 193]. Experiment and experience suggest that the limit on the geometric uncertainty of field edges and all collimation devices should be no more than 4 mm (1 SD), relative to the edges of the PTV[136, 154, 193].

It is important to note that these refer to the total acceptable uncertainties from all components contributing to the radiotherapy treatment. Therefore, the acceptable uncertainty on any individual component should be significantly less. A geometrical accuracy of ±1-2 mm and accuracy of dose calculation or measurement of ±1% are therefore generally aimed towards, both of which are at the limits of what is technically achievable[4, 9, 136, 199].

Implicit in this process is that the CTV→PTV (and OAR→PRV) margins are appropriate[120, 236, 288]: if too small, then the tumour will always be significantly underdosed, resulting in poor tumour control, regardless of how accurately and precisely the dose is delivered; if too large, then far larger volumes of healthy tissue may be receiving a high dose than is necessary. In particular, margins being too large may make it more difficult to optimise the plan to achieve adequate target coverage whilst avoiding sensitive structures, and opportunities for increasing the dose to the target are limited by the significant volumes of healthy tissue also being treated.

It is strongly recommended that all centres regularly review the margins in use and work to improve these as necessary[235, 284, 288]. This is achievable by analysing imaging data acquired during the treatment of many patients, as described below. Although a computationally intensive task this is made relatively straightforward by it being possible to automatically extract data from contemporary RV systems[289].

Dose escalation, and hence better tumour control[65, 162, 299], is possible by reducing the margins added to the CTV to account for setup errors and patient motion. Full advantage from advanced techniques, such as intensity modulated radiotherapy (IMRT) can only be realised in a particular centre if

margins are both appropriate and optimised for the individual treatment site, taking consideration of the facilities available at that centre[133, 155].

1.5. Imaging in Radiotherapy

Imaging is used extensively to support the planning and delivery of radiotherapy[133, 283, 287, 288]. However, before a patient attends the radiotherapy clinic it is likely a significant number of imaging studies will already have been acquired during the diagnosis and staging of the disease. Furthermore, once a patient's treatment is complete, additional imaging will be performed at regular intervals to determine the success of the therapy and verify there is no recurrence. Imaging therefore plays a key role throughout the whole cancer management process, not just in radiotherapy. Figure 1.3 indicates the imaging modalities which might be employed at each stage of the patient pathway, as well as whether the studies are performed under the auspices of the Radiotherapy or Radiology clinical teams. It is important to note that the clinical perspectives and goals are very different between the two teams[162, 287], and this is discussed further in Section 1.6.

In treatment planning, imaging is primarily used to determine the geometrical location and dimensions of the GTV – or the CTV directly if the gross tumour has already been removed by surgery – then to aid beam placement. To ensure both the external body contour and the shapes and configuration of internal anatomy are the same as they will be during treatment all pre-treatment imaging should be performed with the patient set-up as closely as possible to the treatment position, using a flat couch top and the same immobilisation devices. The visualised beams must mimic the exact geometry and capabilities of the actual treatment machine so that the possible treatment scenarios can be accurately 'simulated'.

Ideally, target localisation is performed with cross-sectional imaging so that the tumour can be delineated on each image slice and a complex 3D model of the target constructed[287]. High specification computer workstations facilitate visualisation of the model in the anatomical context provided by the images, and enable generation of the CTV and PTV by automatically 'growing' the identified volume according to specified margins. Other structures, such as OARs and important anatomical landmarks, are also outlined and margins added to these as necessary.

Once a radiation isocentre is chosen, treatment beams are added and their

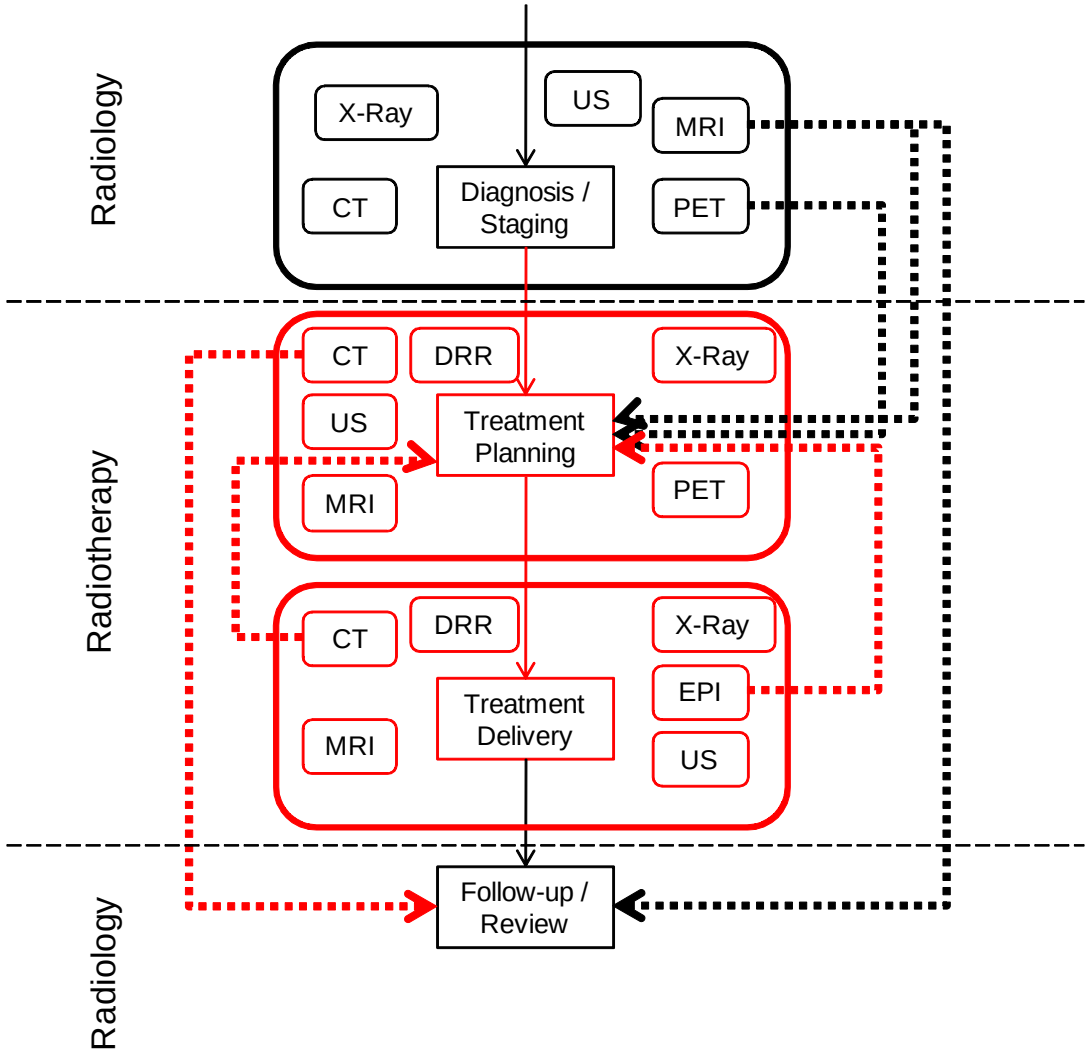


Figure 1.3. – A schematic diagram of the radiotherapy patient pathway, showing the contribution of imaging studies at each stage. Arrows illustrate images acquired at one stage also being used elsewhere. Black indicates studies under the management of Radiology, and red those under the management of Radiotherapy. Note that the imaging modalities identified, and the links between them, are not exhaustive.

collimation and geometry interactively adjusted in the computer to obtain the optimum balance between target coverage and the sparing of healthy tissue. This is achieved through the use of advanced visualisation tools, including the ability to trace field edges through the 3D image dataset. The exact trajectory of a beam can be visualised by generating a planar image of the beam portal superimposed on a superposition of all tissues and structures through which it passes. This *beam's eye view* (BEV)[104] makes it clear which structures are being impinged upon by the beam and enables fine-tuning of the collimation. Furthermore, by ray-tracing through the volume dataset, and applying an appropriate

beam attenuation model, it is possible to predict the X-ray radiograph which would be generated if an imaging X-ray source were positioned at the linac target. These *digitally reconstructed radiographs* (DRRs)[49, 203, 233, 243] can be manipulated to enhance or suppress contrast from different structures and are useful both in treatment planning and delivery verification, as described below.

DRR-like projections can also be generated from volumetric datasets acquired using modalities other than CT. This may be attractive when a particular structure appears with superior contrast on the alternative modality or in specialist cases where CT is not used at all. For example, a number of centres have explored MRI-only planning with some success[49, 176, 212, 213]. In a strict sense, DRRs which are enhanced in some way or which are not generated from CT, so that the resulting image is not intended to exactly mimic a kilovoltage energy radiograph, should be referred to as *digital compose radiographs* (DCRs)[284]. However, because all DRRs/DCRs today tend to involve sophisticated image processing anyway the term DRR has come to refer to any planar digital projection representing a treatment beam geometry.

The process of using tomographic imaging and then assessing beam placement through reconstructions of the 3D data is known as *virtual simulation* because the physical geometry of the treatment machine is being modelled by computer[103, 104, 252]. Combining the visualisation provided by the virtual simulation system with treatment planning software, it is possible to visualise calculated dose distributions in their anatomical context, thus aiding in the optimisation of the overall treatment plan by ensuring adequate target coverage and dose homogeneity, as well as normal tissue avoidance.

X-ray computed tomography (CT) is generally the modality of choice for treatment planning for a number of reasons: it is geometrically robust, provides reasonable soft-tissue contrast at millimetre or sub-millimetre spatial resolution, and there is a well-characterised, reproducible relationship between image pixel values (calibrated in terms of 'Hounsfield Units' or 'CT Numbers') and the tissue electron densities required by the treatment planning system dose calculation algorithms[6, 18, 252]. However, other modalities may also be employed[133, 277], including magnetic resonance imaging (MRI)[172, 177, 277, 300, 340], positron emission tomography (PET)[19, 109, 110, 270] and ultrasound (US)[25, 244]. Table 1.1 summarises recent RCR guidance[287] on the use of imaging to aid treatment planning, and it is clear that both MRI and CT are recommended for the majority of treatment sites. At present, whilst most UK centres make heavy use of CT for virtual simulation and treatment

Tumour Site	CT	MRI	PET	US
Brain	✓	✓	✓	
Head and Neck	✓	✓		
Thyroid	✓			
CNS	✓	✓		
Spinal Cord	✓	✓		
Lung	✓		✓	
Thorax	✓			
Abdomen	✓			
Hepatoma	✓	✓		
Abdomen / Pelvis	✓	✓		
Pelvis	✓	✓		
Lymphoma	✓			
Soft-tissue Sarcomas	✓	✓		
Bone	✓	✓		
Breast	✓	✓		✓

Table 1.1. – Recommended imaging strategies for the treatment planning of different tumour sites. An indicated modality suggests that it might be used with or without contrast media. Multiple modalities for a particular site indicates that images from each of these should be co-registered and interpreted in an image fusion scenario. Adapted from Royal College of Radiologists (RCR) Guidelines[287].

planning, dedicated MRI specifically for radiotherapy applications is not widely accessible[162].

When cross-sectional imaging is unavailable, or where the quality of the DRRs is insufficient, radiographic projection imaging may be utilised instead. Although the images may provide less information overall due to the superposition of 3D information into a single 2D plane, thus making it more difficult to plan complex treatments, simple radiographic imaging has the advantage over virtual simulation that the imaging equipment exactly matches the geometry of the actual treatment machine without requiring computer processing. For this reason, conventional radiographic simulators may be used to verify with the patient present that a virtually simulated plan is technically deliverable, allowing checks of 'basic' technical factors which the computer software may be unaware of, such as whether the linac couch and gantry are in a configuration where they would collide. Although frequently performed in the early days of virtual simulation, increased experience and confidence in the computerised system has led to such plan checking becoming less common. Furthermore,

if the radiographic images disagree with the DRRs provided by virtual simulation, then this may be due simply to patient factors, and it can be difficult to judge which modality is correct. In this situation, the CT scan is still used as the reference dataset for treatment planning because of the 3D information it provides, so if the plan is modified based on measurements made by the conventional simulator there is the risk of introducing a systematic error.

There is no single 'ideal' modality for volume localisation, with each technology providing complementary (and sometimes conflicting) information by nature of its different fundamental operating principles[110, 133, 177, 287, 339]. It can therefore be useful to consider the information provided by different modalities together when determining the GTV / CTV. Modern treatment planning and virtual simulation systems allow image sets from multiple modalities to be loaded simultaneously and co-registered within a common anatomical frame of reference. Structures can be delineated on one modality and then observed superimposed on the others. This *image fusion* harnesses the benefits of each modality whilst overcoming their limitations. For example, CT provides geometrically robust information on the body outline and bony anatomy, but relatively poor soft tissue contrast; MRI provides superior soft-tissue contrast but with less definition of bones and reduced geometrical accuracy. With the widespread availability of digital images through networked hospital information systems, image fusion is becoming increasingly common, particularly between CT and MRI, CT and PET and even between multiple CT scans acquired at different time points or using different contrast agents[177, 244, 277, 304]. Because of the fundamental differences between modalities, and the complex relationship between them when used in image fusion, specialist training and broad experience in image interpretation are essential if the full benefit from each – and the synergy from using multiple modalities together – is to be achieved[287, 288].

During treatment, imaging is used to verify the patient is set-up in the same way at each fraction and that the correct volume is being irradiated throughout. Traditionally, this *verification imaging* was undertaken using X-ray film positioned beyond the patient. However, today this has almost completely been superseded by electronic portal imaging devices (EPIDs), digital detectors which provide superior images for less dose and almost in real-time[181, 330]. Indeed, imaging is completely digital in the new Oxford Cancer Centre and films are used only infrequently in Edinburgh. Film will therefore only be considered superficially in this thesis.

Set-up accuracy is evaluated by comparing on-treatment images against reference images prepared during treatment simulation for the same geometrical projection – either a planar radiograph or, more commonly in contemporary radiotherapy, a DRR[284]. If a set-up error is detected, and the patient set-up is corrected based on the magnitude calculated, additional images may be acquired to verify the final position is correct.

Radiographer and clinician training in image interpretation tends to focus on cross-sectional imaging and projections of coronal or sagittal planes. Furthermore, assessing set-up corrections based on orthogonal image planes is inherently more accurate and robust than with the angled fields frequently utilised for treatment delivery. It is therefore becoming increasingly common that dedicated imaging fields are added, separate from the treatment fields, specifically to facilitate management of set-up errors[284].

In general, set-up errors are categorised as either 'random' or 'systematic'[120, 235]. Random errors are those which vary day to day throughout a patient treatment, and tend to be caused by patient-specific factors such as alignment tattoos being in slightly different positions on the skin surface relative to internal anatomy, or the volume of air in the rectum being different at each fraction. They cannot be predicted in advance, but can be corrected for if identified by imaging at the start of a fraction. Systematic errors result from differences in patient set-up, or the internal configuration of anatomy, between simulation and the treatment machine. For example, the alignment of the isocentre lasers may be slightly different between the CT scanner and treatment machine, or a lung tumour may have been imaged at one extreme of its range of motion due to the CT scan capturing a snapshot of the breathing cycle. Systematic errors can be identified by calculating the average set-up error over the first few fractions of treatment, and then a correction applied for all remaining fractions. Although random and systematic errors are a convenient way of modelling set-up uncertainties, they are a simplification of the overall picture, which also includes factors such as drifting trends (such as a patient losing weight, or lasers gradually moving out of alignment), patients moving suddenly because they are uncomfortable, and involuntary internal motion[236].

Common set-up correction protocols may include portal imaging for the first three fractions, enabling identification and correction of any systematic error, then weekly imaging thereafter, to identify any drift. Regardless of how advanced a model of set-up errors is applied, clinical staff tend to know from experience whether they will have difficulty setting up particular patients

accurately, and the actual imaging protocol is often tailored accordingly.

By analysing the set-up errors for large cohorts of patients it is possible to identify the systematic and random errors for each type of treatment performed in a particular centre[235, 284]. It is strongly encouraged[133, 154, 288, 330] that these are calculated and reviewed on a regular basis and this is now possible by automatically processing data extracted from the RV or associated systems[289]. CTV→PTV (and OAR→PRV) margins can be calculated to account for the set-up errors for different treatment sites, ensuring consistent coverage (or avoidance) of the volumes[120, 236]. The better the set-up correction protocol, the smaller the margins, the less healthy tissue treated and thus the greater the scope for dose escalation and improved tumour control.

Although electronic portal imaging (EPI) is by far the most common form of imaging used for set-up verification, integrated kilovoltage imaging systems and the whole range of cross-sectional modalities are now becoming available. All linac manufacturers now include options for incorporating kilovoltage energy X-ray imaging[160, 173], X-ray CT[56, 192] or cone-beam CT (CBCT)[161, 337, 358] into their treatment solutions, including using the megavoltage treatment beam[273]. Ultrasound[195, 196] and even MRI[279, 280] are also becoming viable options. Utilising tomographic modalities yields a number of advantages. Firstly, megavoltage portal images are inherently poor contrast, so that only bony anatomy can reliably be identified whereas the target volumes tend to be soft-tissue[122, 123]. All tomographic modalities offer superior soft-tissue contrast so provide improved visualisation of the target and avoidance volumes. Furthermore, tomographic images acquired during treatment can be registered and fused with those used for treatment planning, thus enabling a full 3D analysis of patient set-up to be performed. In the most advanced setting, dose calculations can be performed and the treatment plan adjusted for the particular day's set-up. Advanced use of verification imaging is known as *image guided radiotherapy* (IGRT) and active modification of the treatment plan as *adaptive therapy*[155, 250].

Despite the clear advantages from cross-sectional imaging, portal imaging remains the only method of visualising the exact volume through which the treatment beam has passed, being formed by the interaction of the treatment beam itself with the patient. (This is true also for the Tomotherapy system[211], where the equivalent of a portal image is the megavoltage sinogram acquired during treatment, even though this is most easily visualised as a reconstructed CT scan.) An additional benefit of EPIDs is that they are also capable of

dosimetry applications[89]. They can be utilised before a patient's treatment to verify the accuracy of the fluence delivered by IMRT fields[92], and an emerging application is to employ them for *in vivo* dosimetry during treatment, either directly for 2D measurements[231], or – in combination with a CT scan – for full 3D verification dosimetry[90]. EPI will therefore always remain an important modality for set-up verification[285]. EPIDs can also be applied to a range of other physics applications, including equipment output and performance checks[209].

Imaging in radiotherapy is a rapidly advancing field. Emerging applications include using 4D imaging to monitor motion of internal volumes over time. This can be utilised to determine the trajectory of volumes prior to treatment, enabling asymmetric, patient-specific margins to be chosen. A further step is to use this information to develop 'gated' treatments. Together with other measurement devices, imaging may be employed to monitor patient motion during treatment (usually due to the breathing cycle) and the beam turned on only when the target is determined to be within a tightly defined spatial envelope. Advanced delivery technologies allow the collimation to be dynamically adjusted, or the linac head to move, whilst the target is tracked by imaging in real-time[20, 207, 341].

Aside from planning and verification, functional imaging can be utilised to monitor the patient's immediate response to treatment. For example, regular PET scans throughout a fractionation course can be used to identify whether the tumour is shrinking, whether previously unidentified or dormant focii become more active, or whether parts of a tumour are responding differently to others[22, 109, 110]. Like with adaptive therapies, the treatment plan can then be tailored to the specifics of the individual patient.

All conformal radiotherapy relies on imaging to some extent, with the importance of imaging increasing the more advanced the treatment. Whereas imaging may traditionally have been considered an adjunct to the core treatment planning and delivery process, it is now an integral component without which the most advanced treatments could not proceed. Most departments implement adaptive strategies to some extent, with patients being replanned as necessary. As more advanced techniques are implemented the distinctions between pre-treatment imaging, treatment planning and treatment delivery become increasingly blurred.

In a modern hospital, images may be stored in a centralised Picture Archive and Communication System (PACS) which enables studies to be sent via com-

puter networks between different departments. Images acquired for one purpose may therefore be viewed elsewhere to aid decision making for other applications. Increasingly, this functionality is being harnessed in radiotherapy, with images acquired for diagnosis or staging now being used to guide radiotherapy planning[128, 162, 180]. During follow-up, images taken during radiotherapy might also be used as a baseline to evaluate whether a tumour volume has changed size or determine whether there is recurrence of disease. Figure 1.3 illustrates where images might be used for additional purposes. For example, a diagnostic MRI scan or a staging PET/CT scan may be fused with CT in the treatment planning system to aid target localisation. Later, cone-beam CT scans acquired for set-up verification might be used to recalculate a treatment plan, or EPIs fed back into the planning system to help optimise DRRs. The CT scan acquired at the start of radiotherapy planning might also be compared with a 3 month CT scan during a regular diagnostic follow-up appointment. These examples are by no means exhaustive and modern PACS frameworks open up a whole range of opportunities.

As discussed previously, all images used in radiotherapy must be considered in the same geometrical and clinical contexts. However, images taken for diagnostic applications are likely to be acquired using settings and a patient set-up optimised for those purposes, and which might not be optimal for radiotherapy. Using images not specifically acquired for radiotherapy applications therefore introduces significant clinical and technical challenges. When assessing the technical performance of imaging modalities used in any stage of the cancer management process it is important that a consistent methodology is applied throughout, allowing comparisons to be made both between instances of the same modality (such as the CT scanner for diagnosis and that for radiotherapy planning) and between different modalities (such as MRI and CT).

1.6. Optimisation

European[83] and UK[125, 283] legislation requires that all medical exposures involving ionising radiation are *justified* in that the risk associated with the radiation dose is balanced against the benefit to the patient. There is a requirement that doses to healthy tissues are kept "as low as reasonably practicable" and, in radiotherapy specifically, that the dose to each target volume is individually planned. Optimisation of the radiotherapy treatment to achieve an acceptable balance between risk and benefit is a complex process and a key role of the

medical physicist[283, 363].

In a general sense, optimisation is achieving the greatest level of tumour control at an acceptable level of normal tissue complications. On a practical level, this refers to ensuring all the steps which contribute towards the final treatment delivery are individually optimised and balanced against each other, including treatment simulation, treatment planning, treatment delivery and the implementation of set-up verification and correction protocols. Because uncertainties, errors or compromises in any of these can have a negative impact upon the effectiveness of the overall treatment, or can increase the likelihood of short-term side-effects or the long-term induction of secondary cancers, optimisation in radiotherapy requires a systemic approach[283]. Imaging has an important role in each stage of the radiotherapy process, so must be considered as an integral part of this.

Traditionally, optimisation in medical imaging is concerned with balancing image quality with the patient dose burden[222] – either the radiation absorbed dose for modalities involving ionising radiation, or the equivalent for non-ionising modalities[141]. The general maxim is that radiation doses should be kept as low as possible, whilst still being able to yield images suitable for the diagnostic application, with the ideal being that doses should be reduced until images are only just 'good enough' for the intended application. There are also specific legislative requirements intended to enforce this[126], including preventing the imaging radiation beam from extending beyond the sensitive area of the detector, mandating that the quantity of radiation delivered during imaging is accurately measured and ensuring there is active review of patient doses, including the comparison against national reference levels[125].

However, whilst these principles are still relevant, in radiotherapy the imaging is not an end in itself, but rather a constituent part of the overall radiotherapy process[250]. Given that a large dose is intentionally being delivered as part of the treatment anyway, and that imaging doses may be caused by a range of modalities underpinned by different physics, the dose from radiotherapy imaging can be difficult to meaningfully quantify and often difficult to calculate. On a more fundamental level, the task being performed in radiotherapy is considerably different to that in diagnostic imaging. Whereas in diagnostic imaging the goal is to determine whether an abnormality is present or not, or whether it has changed in size, in radiotherapy it is already known that a tumour or lesion exists, with the task being more concerned with accurate geometrical localisation, the determination of physical properties, and co-registrations with and

comparisons against previously acquired images. This is further complicated by the task being different depending upon when images are acquired throughout the radiotherapy treatment pathway. Finally, images in radiotherapy tend never to be considered in isolation. Images from different modalities and different stages in the radiotherapy process are regularly interpreted alongside each other, allowing the continuing fine-tuning and optimisation of a particular patient's treatment. In the widest sense, images from many patients may be considered in the same geometrical and clinical contexts to judge the overall performance of a particular department.

Optimisation of an imaging modality, in terms of balancing dose and image quality, is not the same as optimisation of a treatment[250]. Indeed, there are situations where even a poorly performing imaging modality, in terms of a large dose being required to produce suitable images, may still result in an overall reduction of dose to healthy structures and improved conformation around the target volume. For this reason, dose reference levels have not been introduced for radiotherapy imaging as they have for diagnostic imaging. From experience, the tendency in the radiotherapy community is to either be extremely concerned about imaging doses, such that new techniques such as IGRT are introduced only very slowly (even though a cost-benefit analysis would almost definitely indicate a net dose-saving) or else doses are almost completely ignored (in the case of an ageing and deteriorating portal imaging device). Both scenarios are hard to justify in terms of 'best practice' yet they are difficult to challenge in terms of the overall effectiveness of the contribution of imaging to the optimisation of the treatment. One strategy for radiotherapy imaging optimisation is to consistently quantify the performance of each imaging system involved, then map this to its effect on the optimisation of the overall radiotherapy process.

Ongoing optimisation requires the implementation of a formal programme of Quality Assurance (QA)[125, 126, 283], consisting of regular quality control (QC) checks, planned preventative maintenance (PPM)[154, 193] and the recording and review of these. QA ensures a given system continues to perform satisfactorily over time, enables the identification of performance trends – usually in terms of deterioration over time – and ensures interventions such as repairs or recalibrations can be performed in a timely fashion with the minimum of clinical impact.

QA programmes in radiotherapy typically consists of daily QC checks of critical components, followed by more thorough tests at less regular intervals. Radiotherapy treatment machines and imaging devices are sophisticated equip-

ment which requires considerable expertise and time to test thoroughly. The goal of the frequent QC tests is to quickly ensure key functional components are operating safely within tolerance. Devising a QC test requires a balance between the level of detail and sophistication of the test, and how quick and easy it is to perform. A complicated test takes longer and is more likely to be performed incorrectly, resulting in either an important result being missed or a time-consuming detailed investigation being required. Lengthy tests are also more likely to be postponed or omitted due to the pressures of the clinical day. Therefore, frequent QC checks should be straightforward and quick, yielding unambiguous pass / fail results, so that they can be undertaken by any staff group with appropriate basic training at the start of a busy clinical day. A QC check should also be highly sensitive, so that a single test can rapidly highlight problems across a wide range of components of the overall system.

If a QC test fails there may be insufficient information to determine the cause of the problem, or even the sub-system responsible for the failure, and a more in-depth, and time-consuming, investigation will be required. However, the goal of determining whether the device is performing safely and satisfactorily will have been met. For example, calculating the mean and standard deviation in the Hounsfield Units of the pixel values within a region of interest on a CT scan can be performed in seconds using the software integrated into the control consoles of contemporary scanners, yet this is a very powerful test which assesses the X-ray generation and detection systems, image reconstruction algorithms and image viewing software[252]. Anomalous results could indicate a problem in any of these systems, but further testing is required to localise the problem further. Ideally, additional information is generated during a QC check which can be analysed in more detail later if required. Depending upon the choice of test phantom, this is often possible with imaging tests.

Radiotherapy imaging is underpinned by the computer assisted visualisation of digital images, and the ability to manipulate these as necessary and transfer them between systems. It is therefore essential that any QA programme includes the computer systems themselves, including the performance of the software, integrity of the storage / retrieval system and the final presentation of images on 'soft-copy' display devices. Guidelines [240, 287] suggest that radiotherapy display devices should be of comparable specification to those in diagnostic imaging, and therefore imply they should be subject to similar QA processes[305], but in practice this is rarely considered. Although display devices undoubtedly influence the presentation of the images on which key

decisions are made, including volume delineation and whether a set-up correction should be applied, the clinical impact in radiotherapy of a degraded display device is not understood and must be explored further[297].

Optimisation requires input from all professional groups involved in the patient pathway because each contributes to the overall treatment[283]. Whether a particular treatment is optimal, and whether an exposure is justified, are in the end clinical decisions.

1.7. Image Quality

Image quality is difficult to define in an absolute sense[141]. A 'good quality' image is one which is suitable for its intended clinical purpose[221, 317, 318]. As described above, an image optimised for a radiotherapy application may not be appropriate for use in diagnosis, and vice-versa. Indeed, images acquired during one stage of the radiotherapy process may not automatically be suitable for use at another. For example, a CBCT scan taken during treatment to verify patient set-up may not provide sufficient soft-tissue contrast for volume delineation, or accurate electron density information for treatment planning. In assessing image quality it is therefore essential to consider the fundamental characteristics important to the task being performed. To date, the literature discussing image quality tends to focus on diagnostic imaging, considering radiotherapy applications in only a few specific cases (see, for example[60, 210, 381]). There is only sparse literature covering the overall contribution of imaging to the whole radiotherapy process or the optimisation of radiotherapy imaging in general[250].

In radiotherapy, all images are considered in the same context, and usually influence the sequential progression of the patient through the radiotherapy pathway. The technical quality of an image acquired early in the pathway limits the potential contribution of images taken later in the process. The overall effectiveness of imaging thus tends to be limited by the weakest link in the chain[288]. For example, regardless of how good an EPI acquired for set-up verification is, the effectiveness of the image is limited by the accuracy of the DRR against which it will be compared. However, despite the interdependence of radiotherapy imaging modalities, the performance evaluation and optimisation of each modality still tends to be considered in isolation. The traditional distinction between pre-treatment and verification imaging also still tends to apply, even though the same technologies are now used for

both. All this is in contrast to diagnostic imaging, where there is less clinical inter-dependence between modalities because each investigation contributes an independent piece of information to the construction of an overall diagnostic 'big picture', yet efforts are still made to apply equivalent evaluation methods across modalities.

Ideally, the assessment of image quality involves the calculation of quantitative metrics: i.e. an unambiguous numerical result is produced describing how well images acquired perform under controlled test conditions. However, a fully quantitative evaluation requires the ability to extract the raw pixel values of the image and process these by computer software. Until recently, for some modalities it has been difficult to gain access to the pixel values, including determining the intrinsic processing which may already have been applied to them before extraction from the system. Furthermore, unless a streamlined analysis framework is in place, the processing of the data can be time-consuming, making quantitative analysis inappropriate for regular QC checks. Finally, it is only recently that the final image presented by imaging modalities is completely digital. In older analog modalities an image based on discrete pixels did not exist and it was considerably more difficult to obtain a digitised image for processing, if at all.

Therefore, in both diagnostic and radiotherapy imaging, qualitative methods are currently the norm, with image quality being assessed by visual inspection of images of test phantoms, with the human observer deciding whether or not a particular feature or artefact is visible. By involving more than one observer, and through careful design of the test, these methods become semi-quantitative, yielding results which are reliable within identifiable probability limits. Qualitative and semi-quantitative methods have the advantages of being quick and easy to perform, and that they evaluate the performance of the whole imaging chain, including the presentation of the final image. On the other hand, quantitative methods enable an unambiguous performance assessment of the individual components of the imaging chain, and tests can be designed to consider multiple components working together. Assessment of the presented image is difficult using quantitative techniques, but is possible using specialist equipment. Various research groups have worked to develop observer models which link qualitative methods with quantitative metrics[44, 50, 141, 218, 334]. Although these are successful to some extent, they all depend heavily on the definition of the task being performed. Only two relate specifically to radiotherapy[295, 324].

At present, the methods used to measure image quality in radiotherapy are often modality specific and linked to the historical development and implementation of the particular modality. Techniques are not always compatible between modalities, preventing meaningful comparisons in performance from being made and hindering overall optimisation. Worse, when methods are extended between modalities, they are sometimes implemented in such a way as to be incompatible with equivalent methods applied to the same modalities in diagnostic imaging, thus establishing conflicting baselines between the disciplines and again making optimisation difficult[296].

This thesis describes an approach to the objective and quantitative assessment of image quality which can be applied to all radiotherapy imaging modalities used at any stage of the radiotherapy process and thus aid in treatment optimisation. The same fundamental methodology is applied each time, whilst ensuring performance characteristics important to the specific clinical task are considered. This has the advantage of both simplifying and homogenising the assessment process whilst enabling meaningful comparisons of image quality to be made between the modalities and pathway stages. It thus enables clinically relevant decisions on acceptable levels of performance to be made during commissioning and for quality assurance purposes. Taking the example given above, it would enable the optimum resolution and noise characteristics of EPIDs to be balanced in line with the quality of DRRs generated during virtual simulation. The new framework is also applicable in diagnostic imaging, and to emerging radiotherapy imaging modalities.

1.8. Patient Doses from Imaging

Current radiation protection frameworks are based on the linear no-threshold model of cancer induction, in which even the smallest dose has an associated risk and the level of risk is proportional to dose[137]. This leads to the requirement that any medical exposure involving ionising radiation be clinically justified[125], as outlined in section 1.6 above.

Effective dose[137, 159] is a concept which takes into account both the radiosensitivity of tissues within the irradiated volume and the biological effectiveness of the type of radiation involved. It is widely used as a measure of overall risk in diagnostic imaging. Although recent guidelines[284] suggest the use of effective dose when justifying and optimising radiotherapy imaging protocols, this is extremely problematic for a number of reasons. Firstly, effect-

ive dose represents the stochastic risk of death by cancer induction over the whole lifetime of a healthy patient[137, 138, 159]. It does not account for the drastically reduced life-expectancy of many patients undergoing radiotherapy, the upwardly biased age distribution of cancer patients, nor, in this context, the likelihood of cancer induction from the dose to healthy structures resulting from the actual treatment beams, which is significant[51, 68, 112–115, 250]. In addition, it does not consider the possibility of a reduction in dose to healthy tissue when imaging is used to improve targetting and thus aid in reducing treatment margins[267], or the additional tumour control achievable by dose escalation to the target made possible by this[191, 331]. Furthermore, by intent the PTV is being treated with a deterministic, cell-killing dose, so including any tissues within the PTV in a stochastic risk calculation is meaningless. Finally, the whole concept of effective dose[38, 220] and the validity of the linear no-threshold model are both being questioned by some groups[64, 95, 255], and indeed the most recent ICRP guidelines suggest effective dose may not in itself be a reliable indicator of absolute risk[52, 69, 378].

Despite its limitations, it is still attractive to apply effective dose to radiotherapy imaging because it is a well established quantity within the radiation protection community and there is considerable experience of utilising it in optimisation exercises in diagnostic radiology. In line with the approach taken in diagnostic radiology, from a purely imaging perspective, it can be argued that the concomitant effective dose from radiotherapy imaging exposures should be as low as reasonably achievable whilst providing sufficient information for the clinical task. When formulating an imaging strategy for a particular application – whether this involves selecting from different modalities, choosing from various implementations of the same modality or just modifying acquisition settings – then with all other treatment-related considerations being equal, the lowest dose option is preferable. This is useful in evaluating departmental imaging protocols, commissioning new equipment and when comparing the performance of one department against another. However, there is concern that greater net benefit may be gained by focusing the effort required to drive down radiotherapy imaging doses elsewhere in the radiotherapy process, such as ensuring patient set-up correction protocols and margin calculations are consistent with whatever imaging is utilised[284]. As discussed in section 1.6, radiotherapy is a multi-faceted process and optimisation of imaging involves balancing many considerations against each other, rather than just imaging dose and quality[250]. Also, advanced image guidance equipment has only

recently become widely available so that operators and scientists are not yet confident in the minimum image quality and thus minimum dose required to fulfil particular tasks. There is therefore a tendency to err on the side of 'caution' by being slow at bringing imaging doses down so that the overall clinical application is uncompromised.

Furthermore, because radiotherapy beam geometries are intended to provide an optimum treatment, whereas diagnostic geometries are geared towards optimising image quality, the magnitudes of effective doses and the implications of these may not be directly comparable between disciplines. For example, in diagnostic radiographic imaging the patient dose may be minimised by ensuring the X-ray source and detector are reasonably close to each other, with current European guidelines[47] indicating a standard source-to-detector distance (SDD) for certain projections of 115 cm and other researchers[37, 272] suggesting 100 – 130 cm. However, when imaging a radiotherapy treatment geometry the standard source-to-axis distance (SAD) of 100 cm limits how close the detector can be to the radiation source. To achieve the same fluence and dose at the detector may therefore require technique settings which result in a higher patient dose, yet this does not not necessarily mean that the radiotherapy dose is unduly high. Even though the image quality requirements for diagnosis are generally more demanding it is often the case that radiotherapy imaging doses may unavoidably need to be higher.

Another complication is that effective dose is usually calculated through a combination of experimental measurement and the modelling of dose deposition within the sensitive organs of standard reference phantoms[139, 376]. Although this is accepted practice in diagnostic radiology, where it is not always possible to accurately model a real patient in 3D and optimisation exercises tend to be based on reducing the dose burden for whole patient populations, this is not the normal practice for radiotherapy applications, where routine CT scanning and detailed modelling of doses to target volumes and organs at risk is the norm in treatment planning[71, 162]. Indeed, recent ICRP guidelines[378] suggest tailoring the calculation of risk from an exposure to the details of the individual concerned.

A variety of approaches has been adopted when applying the effective dose concept to radiotherapy imaging. Some workers apply the standard methods directly, with no modification[171, 250, 312, 313], whilst others apply these to CT scans of real patients rather than the reference phantoms[351]. Recognising that the contribution of imaging to the total dose to the target volume during

a radiotherapy treatment is negligible, and that the deterministic cell-killing by intent there voids any stochastic risk calculation, more advanced modelling involves removing the target volume from the effective dose calculation and interpreting the result as a measure of relative risk, rather than absolute likelihood of mortality[113, 251]. This approach may be extended to consider the relative contribution of concomitant imaging doses to individual risk organs as part of the overall treatment[5]. In its simplest form, this may involve measuring or modelling dose deposition in a limited number of sensitive organs[158, 328] and may employ visualisation tools already familiar in radiotherapy treatment planning to help interpret the results, such as dose-volume histograms (DVHs) which indicate the fraction of an organ receiving a particular dose level[71]. More detailed studies involve calculating doses to organs from the primary treatment beams, scatter and leakage from the linac head, then comparing these with those purely from imaging, allowing not just a whole body effective dose but the relative risk to individual organs to be considered. Recent work by Harrison *et al.* has shown that contemporary imaging regimes may contribute between 5 – 25% of the overall treatment dose to particular organs[114, 115].

Most recently, there has been a move away from effective dose to a rather more detailed consideration of the ratio of concomitant to therapeutic dose as a function of distance from the target volume. Generally, imaging fields are slightly larger than the target so that the region of dose delivery can be visualised in its surrounding anatomical context. Therefore, the concomitant / therapeutic dose ratio tends to be highest in the region immediately outside the target volume[112]. This is interesting because it is this region that receives a dose due to scatter and leakage which is relatively high (> 1 Gy) but which is insufficient to reduce the likelihood of secondary cancer induction through sterilisation. There is evidence that induction is most likely here[82] and therefore additional care must be taken to keep concomitant doses in this region as low as reasonably achievable. This is a pragmatic approach which concentrates on potentially the most risky of imaging doses and is achievable using existing dose modelling tools.

Imaging modalities utilised in contemporary radiotherapy employ a range of beam delivery and acquisition technologies which result in exposures with fundamentally different dose distributions[250]. It is therefore difficult to generalise imaging dose assessment and reporting. Some groups are working towards integrating the 3D dose distribution from imaging into the radiotherapy treatment planning process[7, 70, 72], allowing a clinician to fully understand the

contribution of imaging doses to any individual patient's treatment. This may be especially relevant for materials such as bone where the dose deposited by kilovoltage imaging may be particularly high due to increased photoelectric absorption[71]. However, comprehensive implementation of this approach requires knowledge of the full imaging regimen beforehand, which may not be available at the outset of treatment.

It is clear that there is still considerable debate surrounding how to quantify and account for concomitant doses from radiotherapy imaging. However, whichever approach is adopted it is important this is performed consistently across contributing modalities and that the details of the calculation process are discussed and agreed. Otherwise, there is a real danger that presented imaging doses may be misinterpreted and thus hinder optimisation and potentially inhibit the adoption of new modalities which might otherwise provide substantial benefit. This may seem straightforward, but whereas the quantification of the therapeutic radiation dose delivered during treatment is well established and understood, methods for measuring or modelling absorbed dose from imaging exposures have not yet been developed for all modalities, making it extremely difficult to reliably quantify patient dose. For example, in cone-beam CT there are at least three fundamentally different ways of calculating absorbed dose, all yielding results with the same dose units[158, 201, 328, 377]!

Patient doses from imaging is an active research field and a full consideration is beyond the scope of this thesis. Although this work examines the technical relationship between dose and image quality for particular imaging systems, it will not consider patient doses themselves in detail, with it being implicit that a reduced dose to the detector also results in a reduced dose to the patient. Methods of dose assessment for individual modalities are introduced alongside the modality descriptions in the following chapters.

1.9. Goals of This Work

This work aims to provide a framework within which the image quality of any modality can be reliably and objectively quantified, such that the complex balancing act between all competing contributions – of which image quality is only one – can be performed. The actual optimisation of each modality, as part of the overall treatment optimisation, is beyond the scope of this work.

Methods employed for performance assessment in both radiotherapy and diagnostic imaging are critically evaluated, with the goal being to apply the

state-of-the-art in diagnostic image evaluation to radiotherapy, whilst taking into account specific requirements important for radiotherapy. A uniform approach is taken across modalities and disciplines, enabling meaningful comparisons in performance to be made between modalities employed in different stages of the radiotherapy process and between equivalent modalities utilised in the diagnostic world. Concrete illustrations are presented for a number of the most commonly encountered modalities, along with specific applications facilitated by the implementation of this methodology.

A software package *IQWorks* has been developed to support this framework by streamlining image analysis through the development of flexible analysis trees. Being concerned with the objective assessment of image quality, IQWorks is directly applicable not just to imaging in radiotherapy but to modalities across the broad spectrum of medical imaging. By providing a uniform, modality-neutral framework within which objective image quality evaluation can be performed, IQWorks provides valuable information to facilitate the more complex process of optimisation, either of the overall treatment process in radiotherapy, or of a specific imaging modality in diagnostic radiology[292, 296]. It has been released as a free, open-source package to encourage the adoption and evolution of objective analysis techniques and has recently been the subject of a successful national scientific meeting organised by the Institute of Physics and Engineering in Medicine in the UK.

1.10. Structure of This Thesis

This chapter introduced the role of imaging in radiotherapy, the concept of optimisation and the need for an objective and modality-neutral assessment of image quality. The characteristics inherent to a 'good quality' image were discussed.

Methods for the assessment of image quality are covered in Chapter 2. Image formation theory is considered from first principles so that fundamental image quality performance indices can be derived. The importance of the imaging task is emphasised, with the applicability and usefulness of these measures in diagnostic and radiotherapy imaging evaluated. A comprehensive approach to the assessment of image quality in radiotherapy is proposed and IQWorks is introduced as a software framework embodying this. Subsequent chapters then consider a range of different modalities and explore how IQWorks can be applied to these, using existing and novel phantoms. The assessment of image

quality is discussed for each modality, and suggestions are given regarding how optimisation can be approached.

Application of the methodology to megavoltage electronic portal imaging for set-up verification and correction is discussed in Chapter 3. Chapter 4 then covers X-ray CT – including cone-beam CT – for virtual simulation, treatment planning, verification imaging and diagnostic imaging. Finally, brief examples of the applicability of the framework to other imaging modalities employed as part of the radiotherapy process, including radiographic projection imaging, DRRs, MRI, Nuclear Medicine and soft-copy display devices, are presented in Chapter 5.

Themes common to all modalities are discussed in Chapter 6, where the successfulness of the framework and IQWorks is evaluated. Emerging and future applications are also considered and overall conclusions are drawn.

IQWorks is discussed in Appendix A as both a software framework and interactive software package for evaluating image quality. The concept of the 'universal phantom' is discussed and the processing of images by IQWorks 'analysis trees' is described. Design decisions and the advanced technical features of the package are also discussed. Algorithms currently implemented by IQWorks for image quality evaluation are then considered in detail in Appendix B. Essentially, Appendices A and B demonstrate the modality-neutral implementation of the fundamental theory detailed in Chapter 2.

Examples of papers presented by the author on this work, modalities to which IQWorks has been applied, and details of how to obtain and run the IQWorks software itself are included in the other appendices.

Chapter 2.

Quantitative Assessment of Image Quality

2.1. Overview

Methods for the performance evaluation of imaging systems have existed ever since any given modality was first applied in radiotherapy. However, there is a tendency for these to be linked to the traditional approaches taken by the early adopters or pioneers of the technology. Whilst appropriate for assessing any given modality in isolation, they are not generally applicable between modalities, thus making inter-modality performance comparisons and the optimisation of the overall radiotherapy imaging pathway difficult. Furthermore, because diagnostic and radiotherapy applications have historically developed in parallel, with little overlap, there are instances where supposedly identical metrics between the two disciplines are actually incompatible.

For example, two phantoms[275, 281] commonly in use in the radiotherapy community for quantitatively assessing the spatial resolution of electronic portal imaging devices (EPIDs) employ Coltman's method[54], as described by Droege and Morin[80], to determine the modulation transfer function (MTF) through measuring the variance in regions of interest (ROIs) on a series of bar patterns of different spatial frequencies. In principle, these phantoms should yield results consistent with those of the edge phantom method well established in diagnostic imaging[308]. However, due to differences in the way in which the MTF curve is normalised between the approaches the results are incompatible, yet these are widely disseminated in the literature without any comment of this (see, for example [21, 232, 239, 308]). This situation appears to stem from a decision taken by the designers of the earlier of the two EPID phantoms[281] and has been perpetuated over time[123], even into newer phantoms designed

to assess kilovoltage radiographic detectors[278, 372], and has resulted in an inconsistent approach between the radiotherapy and diagnostic imaging evaluations of detectors with the same underlying technology. There is no good reason for such a fundamental difference in implementation of the same underlying theory[54, 80], and it hinders modality intercomparisons. Worse, it even prevents intercomparisons between measurements made using different bar-pattern phantoms on the same modality. This is demonstrated in detail in chapter 3 and IQWorks is explored as a tool for implementing a consistent approach across all currently available phantoms and radiographic imaging systems.

Another example of inconsistent approach between radiotherapy and diagnostic imaging is in the uniformity assessment of cone-beam CT (CBCT). Standard practice in diagnostic CT is to place five ROIs on an image of a uniform circular phantom – one at the centre of the field of view (FOV) and the others at the periphery – then a metric of uniformity is calculated from these[151, 156]. Although Varian Medical Systems (Palo Alto, California, USA) follow a similar methodology in the the formal evaluation procedure of their CBCT systems, and employ the same CT phantom, the metric calculated is subtly different[345, 359]. Care must therefore be taken when comparing results calculated by different groups for different CBCT systems, even though the same phantom may have been used in the studies. Furthermore, the Varian procedure is specifically limited to consider only a small field of view, of the order of only 15 cm diameter, whereas in diagnostic imaging the evaluation area may be chosen appropriate to the given application (e.g. a large diameter uniformity phantom and measurement FOV for a large FOV imaging mode). Again, there is no good reason for this difference in approach, especially given that the maximum FOV of the Varian CBCT systems is of the order of 45 cm. As well as hindering intercomparisons between CT modalities this also inhibits optimisation exercises, where the clinically useful FOV may be larger than that included in Varian's analysis scheme. Assessment of CT uniformity is explored in chapter 4.

It is worth noting that discrepancies also exist between performance metrics calculated for the various diagnostic imaging modalities. Taking uniformity assessment as an example again, in magnetic resonance imaging (MRI) the standard approach is to calculate a metric based on the pixel values in horizontal and vertical line profiles across the FOV[197, 198], whereas in nuclear medicine it is normal to apply a smoothing kernel to the image before any

analysis is performed, then different line profile or ROI based techniques are applied[327]. In this way, uniformity metrics calculated in the standard way for Varian CBCT, diagnostic CT, MRI and nuclear medicine modalities will all be different quantities which cannot be used directly for relative performance comparisons. Sometimes the differences in implementation are deliberate – a good metric calculated as part of a consistency test within a routine quality assurance programme should concentrate on weaknesses of the modality and be sensitive to material changes in performance. However, at the same time it is important when optimising performance across a multi-modality imaging chain to be able to apply the same numerical algorithm to different modalities, and IQWorks provides an opportunity for doing this.

Now that high quality imaging is available at all stages of the radiotherapy process, full optimisation of treatment planning and delivery requires optimisation of the contribution from imaging at each stage. This in itself requires an understanding of the limitations and benefits of the different modalities, particularly regarding their influence on the accuracy and precision of dose delivery. Detailed performance evaluation of the different modalities is only possible if they are considered within a common analysis framework. In addition, convergence between diagnostic and radiotherapy imaging, with images originally intended for one discipline now routinely being applied to the other, requires a common approach not just between different modalities, but across all disciplines.

Quantitative assessment of image quality involves calculating numerical metrics to describe an observer's ability to perform a particular task. Inherent to any task is the process by which the observer considers and makes decisions based on the information presented in the image. Regardless of the physics underpinning any particular modality, imaging performance can be described by the same fundamental imaging science[141]. Although this is the case, there has been little effort in the past to apply this theory in a radiotherapy context, or indeed to link the classical metrics widely utilised in radiotherapy back to the underlying theory.

This chapter introduces the signal detection theory (SDT) which underpins the radiotherapy imaging task. Following a discussion of SDT, transfer theory is introduced and the potential of macroscopic or 'large area transfer' metrics explored. Image formation theory is then considered, yielding more descriptive metrics in the spatial-frequency domain. The clinical relevance of these theoretical constructs is then discussed by considering their roles in observer

models.

Throughout this chapter, the image evaluation methods classically employed in radiotherapy are compared with more advanced and universally applicable analyses, with a pragmatic approach being taken to the implementation of these in a clinical radiotherapy environment. Recommendations are formulated regarding concrete, practically realisable image quality metrics for radiotherapy, and a complete list of these is collated in Appendix C. The IQWorks software framework for applying these methods as part of a radiotherapy commissioning or a routine QA programme is presented in chapter A, and the numerical methods for calculating the metrics in chapter B.

2.2. Signal Detection Theory

All images contain information relevant to the task being performed (i.e. the 'signal'), as well as noise which detracts from the task[62, 141]. Because the noise introduces uncertainty into the presentation of the signal there is always a finite probability of making the wrong decision, with the accuracy of decision-making dependent upon the relative characteristics of the signal and noise, as well as the criteria the operator is applying to make the decision. In the discussion which follows, the exact nature of the decision itself is unimportant, but it should be taken that the decision required is appropriate to the clinical task.

Detection of the signal, and the identification of the absence of signal, are both outcomes which belong to probability distributions. Illustrated in Figure 2.1 are the signal ('S') and no signal ('NS') probability distributions for an image acquired by a hypothetical imaging system under particular configuration settings. The vertical 'decision criterion' line indicates the confidence threshold the operator has chosen when interpretting the image, to the right of which everything is perceived (whether correctly or not) as signal, and everything to the left as no signal. From the figure it is evident that the decision may fall into one of four categories:

- the signal is correctly identified, the likelihood of which is given by the area under the blue S curve, to the right of the decision criterion line – the True Positive Fraction (TPF) of results;

- regions of no signal are mistakenly identified as signal, corresponding to the area under the red NS curve beyond the decision criterion line (red +

grey shaded areas in the graph) – the False Positive Fraction (FPF);

- areas of no signal are correctly identified – the True Negative Fraction (TNF);
- the signal is missed, being mistakenly thought to be a negative result – the False Negative Fraction (FNF).

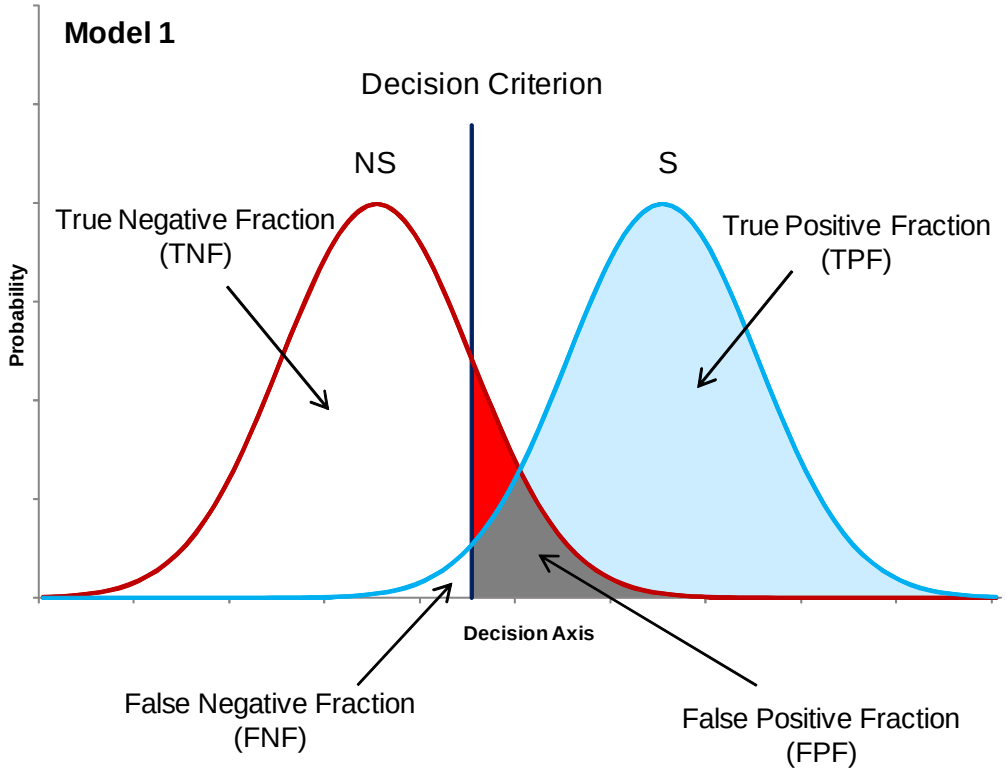


Figure 2.1. – Probability distributions associated with the presence of the signal ('S') and no signal ('NS').

In a clinical scenario, the goal is to maximise the TPF whilst minimising the FPF. When assessing imaging modalities, the TPF is referred to as *sensitivity*, because it relates to the detectability of a signal, whilst $1 - \text{FPF}$ is defined as the *specificity*, describing the reliability of identified positive results[141, 221, 318]. Maximising task performance therefore maps to maximising both sensitivity and specificity, through optimising image acquisition settings and the choice of the decision criterion. The nature of the probability distributions depends upon the imaging system and its acquisition settings, whereas the decision criterion is intrinsic to the operator.

Noise in an image affects the spread of the distributions, whereas signal magnitude influences their separation along the decision axis. If the signal

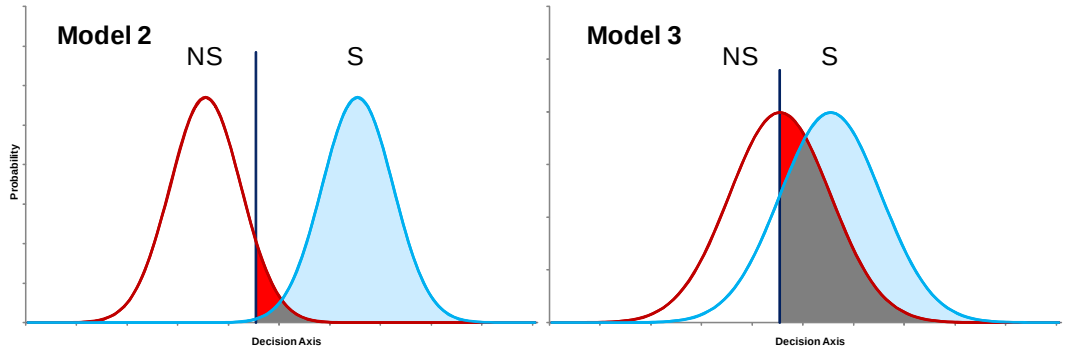


Figure 2.2. – Effect of noise and signal magnitude on decision making. In **Model 2** (left) the noise is less than in Figure 2.1 whilst the signal magnitude is the same. In **Model 3** (right) the signal magnitude is less, whereas the noise is the same.

magnitude is kept the same but the noise level reduced, the curves becomes more 'peaked', so that both sensitivity and specificity improve, as shown in Model 2 in Figure 2.2. Alternatively, if the signal amplitude is reduced, but the noise kept constant, the separation between the distributions decreases and sensitivity and specificity are reduced, as illustrated by Model 3 in the figure. Clearly, if the distributions are well separated, then the task becomes easier, whereas if they overlap the task becomes harder.

For a real system and task, the decision criterion may or may not be a conscious decision of the operator, it may vary depending upon the importance of the task or it may develop as a result of training and experience. However, regardless of how good the operator, the accessibility of the information content is always limited by the nature of the S and NS probability distributions.

For a given pair of distributions, a plot of sensitivity against specificity for all possible decision criteria is known as a receiver-operator characteristic (ROC) curve. The area under the curve is indicative of the best possible performance achievable by the operator, using the imaging device with particular acquisition settings, and is thus a measure of overall image quality[141]. The ROC curves for the three models presented above are plotted in Figure 2.3.

One strategy for optimising performance is therefore to aim to adjust the configuration of an imaging device (including acquisition settings, geometry, etc.) to maximise the area under the ROC curve. This is commonly attempted in diagnostic imaging. However, the disadvantage is that it requires measurement of the ROC curve each time a setting is changed, using the subjective / semi-quantitative tests outlined in Section 1.7. Repeated measurements by different observers are required to achieve statistically meaningful results, making the

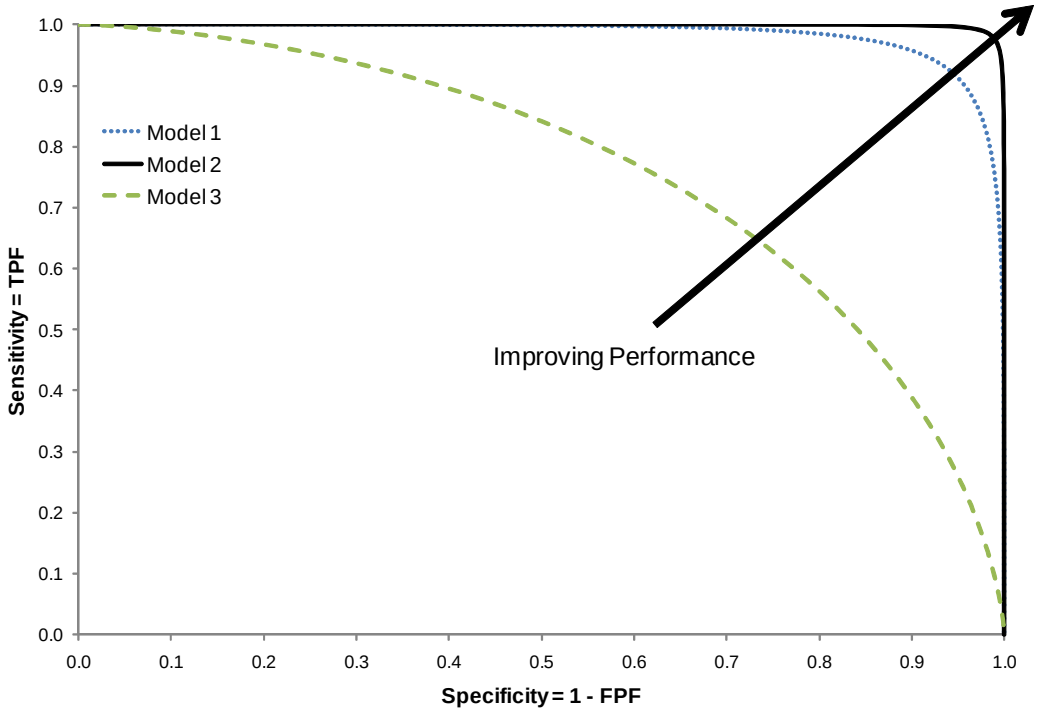


Figure 2.3. – ROC curves for the models in Figures 2.1 and 2.2.

process very time-consuming and labour intensive. Furthermore, care must be taken to ensure the task being performed when interpretting the test objects is representative of the clinical task, yet this can be extremely difficult[189, 334].

An alternative approach is to consider the nature of and relationship between the S and NS probability distributions themselves. A related metric of image quality is the 'discriminability index', d', which describes how easy it is to distinguish which distribution a result belongs to. d' is given by

$$d' = \frac{\Delta}{\sigma} \tag{2.1}$$

where Δ is the separation between the centres of the curves, and σ is the standard deviation of the noise[119]. This analysis assumes the noise is Poisson distributed, which is usually the case for imaging systems relying on random processes, such as the emission and detection of X-ray photons. However, there may also be other sources of noise, potentially belonging to non-Poisson probability distributions, such as those encountered in MRI, US or even introduced by the human visual system. If required, these can be accounted for in the more sophisticated methods described below.

When applying statistical decision theory to imaging science, the manifestation of d' is the signal-to-noise ratio (SNR)[141, 334], which is thus an overall

metric of image quality encompassing all technical elements affecting the performance of the specified task. However, accurately specifying a real clinical task can be extremely difficult (if not impossible), so often a simplified task is chosen to aid in technical optimisation.

Determining the minimum SNR required for an acceptable sensitivity and specificity requires modelling the behaviour of a human operator. Various models, of differing degrees of sophistication and fidelity to the task, are described later in Section 2.7. However, it is generally thought that the SNR must be greater than approximately 3 – 5 to achieve acceptable performance[44, 73, 219, 302]. Once the SNR has been optimised by technical experiment and numerical analysis, a subset of the most promising scenarios can then be put forward for the labour intensive ROC analysis, thereby testing theoretically optimum results using a human observer in a clinical environment[354].

2.3. Transfer Theory

Formation of a medical image consists of three stages:

- image acquisition – the physical interaction of radiation with the object being examined, then the interaction with and detection of the radiation by a detector;
- image storage – digitisation and encoding of the detected image so that it can be stored in computer memory, or written to a storage medium for future retrieval, often with additional information embedded (e.g. patient name, pixel size, etc.) ;
- image presentation – display of the image for interpretation by an observer, usually on a soft-copy device such as a computer monitor or projector, with the data being manipulated or enhanced as required (e.g. windowed, edge enhanced, etc.).

Each of these stages itself involves a number of sub-stages. At each stage, a signal enters as input, is processed, then is output to the subsequent stage. Image quality depends upon how faithfully the information contained in the object being examined is 'transferred' through each stage of this imaging chain. 'Transfer theory' models the passage of information through the system. At each stage the object signal will be subjected to various processes which enhance, distort or degrade the information content.

Information transfer is affected by:

- the system characteristic curve — which relates the input signal at the detector to the pixel values in the image;
- noise — both random and structured;
- spatial resolution — a measure of the 'sharpness' of the image and the size of objects which can be visualised.

Transfer theory is greatly simplified if it can be assumed that the characteristic curve is linear: i.e. the output is directly proportional to the input.

A system is linear if and only if its transfer characteristics are such that an input $h(\mathbf{r})$ at point $\mathbf{r}$ results in an output signal $S\{h(\mathbf{r})\}$, where for any two individual inputs $h_1(\mathbf{r})$ and $h_2(\mathbf{r})$

$$S\{h_1(\mathbf{r}) + h_2(\mathbf{r})\} = S\{h_1(\mathbf{r})\} + S\{h_2(\mathbf{r})\} \tag{2.2}$$

and

$$S\{ah(\mathbf{r})\} = aS\{h(\mathbf{r})\}$$

for any real constant a.

Even though most imaging modalities are generally not linear, it is usually possible to design the experiment or manipulate the data so that they can be considered so. For example, many modalities exhibit linearity over a small range of signal values, so linear transfer theory still applies if the test object being considered is such that the signals present in the image are within this linear range. Alternatively, inherently non-linear systems (such as film-screen systems or computed-radiography) can be 'linearised' by measuring and applying a calibration curve. All the radiotherapy imaging modalities considered in this study are either inherently linear or can be linearised, although sometimes judicial choice of the signal units is important. For example, a CT slice traditionally contains pixel values presented in Hounsfield Units (HUs), which are proportional to linear attenuation coefficient. However, the same CT data loses their linearity if presented instead as an electron density map. Therefore, the first step in performing image quality evaluation is to ensure the system is properly linearised.

Macroscopic, or 'large area', transfer characteristics describe information propagation using generalised metrics and tend to be based in the spatial domain. These are useful for assessing gross performance, and will be considered in section 2.4. A more detailed model of information transfer, operating primar-

ily in the spatial-frequency domain and constructed from the fundamental image formation process is presented in 2.5.

2.4. Large Area Transfer Metrics

Analysing regions of interest (ROIs) in the spatial domain enables gross metrics of image quality to be calculated.

Consider an image with two ROIs, ROI_o covering the object being examined (i.e. the signal) and ROI_n in the background noise, with mean pixel values $\overline{s_o}$ and $\overline{s_n}$, and pixel standard deviations of σ_o and σ_n.

The signal-to-noise ratio is then given by either[62, 197, 217]

$$SNR = \frac{\overline{s_o}}{\sigma_n} \tag{2.3}$$

or

$$SNR = \frac{\overline{s_o}}{\sqrt{\sigma_n^2 + \sigma_o^2}} \tag{2.4}$$

depending upon whether just the noise in the background is being considered, or an average of the noise between the object and background regions. Typical observer models assume that all noise is random and that there is no noise in the signal itself. Under these conditions, one would expect σ_o and σ_n to be very similar, and on average the same.

However, for real images, the noise level in ROI_o tends to be a function of $\overline{s_o}$ because the contrast of the object itself depends on the quanta of radiation contributing to its image and is thus subject to Poisson statistics. If N quanta are detected then the noise $\sigma = \sqrt{N}$. Thus, if the signal s_o is proportional to N_o then the noise $\sigma_o \propto \sqrt{s_o}$, leading to the seemingly counter-intuitive result that the noise level, in absolute terms, increases with larger signals. It can therefore be more pragmatic to calculate an average noise value between the background and signal regions, as in equation 2.4.

Signal detection theory is concerned with the detectability of a signal above the background noise. If the background region itself, in the absence of noise, contains a significant offset signal then the difference in signal between the the object and background regions becomes important. The 'difference' or 'detail' signal-to-noise ratio (dSNR) is given by either[281, 333, 335]

$$dSNR = \frac{C}{\sigma_n} \tag{2.5}$$

or[77, 237]

$$dSNR = \frac{C}{\sqrt{\sigma_n^2 + \sigma_o^2}} \tag{2.6}$$

where C is the 'contrast' between the two regions and is a measure of the difference in signal between them. It is defined variously as simply the difference in signal between two regions[197, 333]

$$C = \overline{s_o} - \overline{s_n} \tag{2.7}$$

the ratio between the signal difference and that of the background only[94, 263]

$$C = \frac{\overline{s_o} - \overline{s_n}}{\overline{s_n}} \tag{2.8}$$

the ratio between the difference and the sum of the average signals in the two regions[67, 237]

$$C = \frac{\overline{s_o} - \overline{s_n}}{\overline{s_o} + \overline{s_n}} \tag{2.9}$$

or the ratio between the difference and overall average signal between the two regions[123, 309]

$$C = 2\left(\frac{\overline{s_o} - \overline{s_n}}{\overline{s_o} + \overline{s_n}}\right) \tag{2.10}$$

Contrast does not by itself directly indicate detectability but rather can be used to determine or verify the characteristic curve of a system.

$dSNR$ is sometimes referred to as the contrast-to-noise ratio (CNR) and is a useful metric for constancy testing or optimising an individual system. However, because it depends upon the system characteristic curve care must be taken when using it to make comparisons between different examples of the same modality. Furthermore, because the contrast between objects also depends upon the nature of the radiation source it is inappropriate to use $dSNR$ for inter-modality performance comparisons.

A measure of the change in contrast between stages of the imaging chain is given by the large-area transfer factor T_C which is the ratio of contrasts measured at the same locations in the output and input images.

$$T_C = \frac{C_{out}}{C_{in}} \tag{2.11}$$

On its own, T_C does not adequately reflect overall image quality because it is possible for large-area contrast to be preserved whilst fine detail is lost, such as in the blurring of small objects or edges. Therefore, structure size and spatial resolution must also be taken into consideration. In addition, because practical measurement of contrast relies on taking statistical averages of pixel values in macroscopic ROIs, T_C is also potentially insensitive to even high levels of either random or structured noise. Characterising large-area contrast transfer across the whole imaging chain requires knowledge of the contrast inherent to the object being imaged, taking into account the interactions with the object of the radiation involved. This can be extremely difficult to determine for some modalities and test objects. It is also questionable whether the results obtained map meaningfully to the clinical sphere, where the objects being imaged are considerably more complex and contain a wide range of signal values and spatial-frequencies. As described in Section 2.5, a consideration in the spatial frequency domain provides a deeper understanding of how the information in the final image reflects that in the original object.

In a similar manner to the signal of interest, noise is also transferred between the stages of the imaging chain. Stochastic noise results in fluctuations in image signal, as a function of either position or time. Noise level can be described by the variance in the expected signal value in a region which, in the absence of noise, would be uniform. Using the nomenclature above, the noise variance is given by

$$\sigma_n^2 = E\left\{|\Delta s_n|^2\right\} \tag{2.12}$$

where $\Delta s_n = s_n - E\{s_n\}$ and the expectation value $E\{s_n\}$ is obtained by averaging the signal value s_n at the same location $\mathbf{r}$ in many repeated images (an ensemble average). However, repeated measurements are often impractical, so a spatial average tends to be used as an estimate instead. A system where the ensemble and spatial averages are equivalent (i.e. $\overline{s_n} = E\{s_n\}$) is known as being as *ergodic*.

Macroscopically, the influence of noise can also be described using the coefficient of variation (CoV):

$$CoV = \frac{\sigma_n}{\overline{s_n}} = \frac{1}{SNR} \tag{2.13}$$

with the intention being that the CoV provides a better indication of the relative degradation in detectability due to noise. However, whilst providing informa-

tion about overall noise level, assessments based on variance measurements are of limited value because they do not describe spatial correlations in the noise, yet two images with the same noise variance may appear very different due to the appearance of structure caused by these. Considering the system in the spatial-frequency domain enables the nature of correlations to be examined and thus provides a fuller characterisation of the noise and its transfer through the system.

If noise is truly due to random processes then the pixel values in a region of uniform input signal should have a constant variance. However, detector issues – such as defective pixels or a spatially varying characteristic curve – will add non-stochastic structure to the noise, although this is often difficult to identify by eye because in-built image processing algorithms may automatically be applied to correct for this. One straightforward, but powerful and sensitive tool for examining local noise characteristics in the spatial domain is the 'variance map'. This is an image in which each pixel value is equal to the variance of a set number of surrounding pixels in the original image. From visual inspection it is immediately clear whether the noise level is changing significantly with position, and in particular whether there are any localised regions of structure. For truly random noise the variance of the variance map should be zero. A quick QC check can be performed by considering an ROI on the variance map itself and verifying the variance in the ROI is below a set tolerance.

Another method of investigating spatial variations is to consider line profiles: tabulating pixel value as a function of distance along a line. Across regions of uniform exposure, a profile will reveal information both about noise characteristics and the nature of the characteristic curve. Alternatively, if the noise and characteristic curve are well understood, profiles will yield information about the radiation source and the object being examined. In a particularly noisy system, the contribution of random noise can be reduced by averaging adjacent line profiles: i.e. taking an area profile.

Profiles can either be plotted and inspected by eye or a variety of metrics of uniformity may be calculated from the pixel values along the profile line. In the equations which follow p_i represents the numerical value of the ith pixel along the profile line when stepping incrementally from the profile start point.

The coefficient of variation along a profile is calculated in the same way as for ROIs:

$$CoV = \frac{\sigma_p}{\overline{p_i}} \tag{2.14}$$

Integral uniformity U_i assesses the maximum deviation from the mean:

$$U_{i(+)} = \frac{p_{max} - p_{min}}{\overline{p_i}} \tag{2.15}$$

and can be divided into positive and negative components:

$$U_{i(+)} = \frac{p_{max} - \overline{p_i}}{\overline{p_i}} \tag{2.16}$$

$$U_{i(-)} = \frac{p_{min} - \overline{p_i}}{\overline{p_i}} \tag{2.17}$$

Another metric, concerned with more local variations, is the differential uniformity U_d which measures the maximum difference between adjacent pixels, and as such is very sensitive to noise spikes and detector artefacts:

$$U_d = \frac{\Delta p_{max}}{\overline{p_i}} \tag{2.18}$$

Depending upon convention, the denominator in equations 2.15 to 2.18 may instead be $(p_{max} + p_{min})$, with the different approaches being analogous to those adopted for contrast assessment in equations 2.8 and 2.9.

Finally, an alternative approach is to identify the proportion of pixels with values outside a set tolerance, e.g. $\pm$ a specified number of standard deviations about the mean value. This is known as the fractional uniformity U_f.

Which metric of uniformity is most applicable will depend upon the task being performed. For QC checks it is the one which most readily indicates a significant change in performance with acceptable sensitivity and specificity. Often, when calculating global uniformity metrics only a specified portion about the centre of the field of view is considered (e.g. 90%) in order to avoid the influence of known edge effects. Larger scale measures of uniformity can also be calculated by applying the metrics above to the pixel values within whole regions of interest.

Discriminability depends not just on the characteristics of contrast and noise transfer but on how detail is preserved. 'Spatial resolution' is a measure of the size of objects which can be observed in an image. In the spatial domain this may be defined as:

- the dimensions of the sensitive elements in the detector, or the distances between the centres of these (i.e. detector element 'size' or 'pitch');
- the dimensions of an individual pixel in the image matrix (i.e. pixel 'size' or 'pitch');

- the smallest object which can be discerned;
- the smallest separation which can exist between two structures before they merge together, becoming no longer distinguishable as separate objects.

Although similar, depending upon the modality and imaging conditions these definitions are not necessarily equivalent. However, they all relate to the discernability of small objects or small gaps between objects.

On a macroscopic scale, resolution may be assessed visually be examining a series of periodic bar patterns, where in each successive pattern the bars become narrower and closer together (i.e. the spatial frequency increases). The 'limiting resolution' is taken as the highest spatial frequency pattern just clearly visible under controlled conditions, and is usually expressed in spatial frequency units (line pairs per unit distance or cycles per unit distance), and sometimes simply as the spacing of the pattern. Often the bar patterns are designed for maximum contrast so that the influence of noise is minimised, and under these conditions the quantity measured is the 'limiting high contrast resolution'.

However, a limitation of this large-area approach is that only the high frequency performance of the system is assessed, whereas there is evidence that low-frequency transfer has an important role to play in the diagnostic task[57]. Although not considered in the literature, it is logical to assume that radiotherapy imaging activities also rely on low-frequency components. For example, in the image registration / fusion task an operator must compare equivalent structures against each other on multi-resolution, multi-modality images, intrinsically performing the comparison on both global and local scales. A full characterisation of spatial resolution therefore requires a more detailed analysis in the spatial frequency domain, rather than simply inspecting repeating patterns in the spatial domain.

Practical experience reveals that the large-area transfer metrics described above are interrelated. Detectability relies upon a complex relationship between different aspects of the imaging system and the object being imaged. Whether an object can be distinguished above background noise depends upon its size, the noise level in the image and the contrast between the object and background. When assessing image quality in an absolute sense it is therefore difficult to determine exactly which metrics to apply. Decisions must be made regarding which definition of SNR, contrast and dSNR should be utilised, and indeed whether dSNR is more appropriate than SNR. However, it is suggested that for a basic constancy check consistency of approach is more important than

accurate modelling of the clinical task — as long as the same, possibly non-optimal, methodology is applied each time then degradation in performance can successfully be identified.

Macroscopic image quality metrics clearly do not provide a complete description of the system, with it being possible to gain a fuller understanding through a more detailed consideration in the spatial frequency domain. However, large-area metrics tend to be intuitive and provide a global perspective on performance which can often be lost when focusing on the technical details. They are also straightforward to calculate and are useful for quick QC checks.

2.5. Image Formation Theory

Throughout the steps of the imaging chain the information contained in the object being examined is mapped into what is visualised as the final image.

At the most fundamental level, the information content of the image is limited by the inherent contrast provided by the interaction of radiation with structures in the object. This is a function of the type of radiation (and hence the modality itself), acquisition settings and the physical characteristics of the object. Regardless of the amount of information originally present in the object, if structures are 'invisible' to the radiation being used to probe it then the information contained in these cannot be transferred to the image and will be lost. One can therefore define the 'best possible' image as that which, free from noise, faithfully contains all information accessible by the radiation of choice. This will never be a 'perfect' image because no radiation is capable of differentiating between all structures.

Emission and detection of electromagnetic radiation — whether ionising or non-ionising — are random processes involving quanta of energy (photons) and as such are subject to Poisson statistics. The 'quantum image' is the 'virtual' snapshot in time of the radiation quanta after interaction with the object. It contains all the information present in the object, limited by the fundamental physics of the interaction processes, and is a degraded sampling of the continuous spatial distribution which is the best possible image. In this image, each individual quantum exists at a discrete point $\mathbf{r}_0$, sampled from a distribution of possible positions $\tilde{\mathbf{r}} = \{\mathbf{r}_i\}$. Its contribution can be described by the Dirac delta function $\delta(\mathbf{r} - \mathbf{r}_0)$, with the overall image $q(\mathbf{r})$ being the superposition of

a number of these:

$$q(\mathbf{r}) = \sum_{i=1}^{N_q} \delta(\mathbf{r} - \mathbf{r}_i) \tag{2.19}$$

Averaging many instances of equation 2.19 yields the expectation function $E\{q(\mathbf{r})\}$ and describes the quantum image in the absence of noise. This is equivalent to the 'best possible' image, degraded only by technical limitations of the radiation source (such as target spot size in an X-ray tube). If the quantum image consists only of a Poisson distribution of quanta then $\tilde{\mathbf{r}}$ is randomly distributed and all contributions are uncorrelated across the image area.

Images are acquired over a finite period of time. Assuming all quanta interacting during a time-window Δt contribute to the final image, the quantum image $q(\mathbf{r}, \Delta t)$ can be considered a time-integrated series of instantaneous samples of $q(\mathbf{r})$:

$$q(\mathbf{r}, \Delta t) = \int_{t=0}^{\Delta t} \left(\sum_{i=1}^{N_q} \delta(\mathbf{r} - \mathbf{r}_{i,t}) \right) dt \tag{2.20}$$

As Δt increases the relative uncertainty due to random noise decreases and the quantum image tends towards the expectation function, as long as there is no motion of the object or radiation source.

$$E\{q(\mathbf{r})\} = \lim_{\Delta t \to \infty} \{q(\mathbf{r}, \Delta t)\} = \int_{t=0}^{\infty} \left(\sum_{i=1}^{N_q} \delta(\mathbf{r} - \mathbf{r}_{i,t}) \right) dt \tag{2.21}$$

In practice, optimisation of the time window involves balancing the information content yielded from a quantum image with superior noise characteristics against the smearing effect of motion artefacts.

Visualisation of the quantum image requires counting of the quanta present at each location. Most modalities detect the presence (or absence) of quanta through absorption or scattering interactions, with the resulting signal being biased by the effect of the energy of each quantum on the interaction mechanism. At this stage, one can consider the 'analogue image' $d(\mathbf{r})$ as the conversion of the quantum image into a continuous spatial distribution of signal values sampled from a continuous range of signal values. It represents the quantum image weighted by the energy response of the detector and degraded by the physics of the detector interaction mechanisms.

In contemporary modalities the analogue image is digitised to enable storage, visualisation and image processing. The 'digital image' $p_{x,y}$ thus represents $d(\mathbf{r})$ at a discrete set of spatial coordinates – usually on a rectilinear grid (n_x, n_y)

– and mapped to a set of permissible discrete values. Generally, for most modalities the formation of the analogue and digital images are complex, multi-stage processes which are often interrelated.

Image quality metrics are outlined below by considering the degradation of the digital image in comparison with the quantum image. For clarity, concepts are outlined for the 1-dimensional case in the spatial domain x and spatial-frequency domain u, but these are immediately applicable for 2- or 3-dimensional datasets. As discussed in Section 2.4 it is assumed that the imaging system is intrinsically linear or has been linearised.

If a photon is detected at position x_0 the response S of the detector to the input impulse is given by the 'point spread function' (PSF):

$$S\left\{\delta\left(x-x_0\right)\right\} = PSF\left(x, x_0\right) \tag{2.22}$$

where $PSF\left(x, x_0\right)$ means 'as a function of x, due to an impulse at x_0'.

By extension of equation 2.19 the complete detector response is the superposition of the many detected impulses

$$S\left(x\right) = \sum_{i=1}^{N} PSF\left(x, x_i\right) \tag{2.23}$$

and, by equation 2.2, in a linear system this image can be considered simply the sum of the contributions from the PSF presented at each impulse location, potentially weighted by an energy factor k_i:

$$S\left\{\delta\left(x-x_1\right)+\delta\left(x-x_2\right)+\ldots+\delta\left(x-x_N\right)\right\} \tag{2.24}$$

$$= k_1 PSF\left(x, x_1\right) + k_2 PSF\left(x, x_2\right) + \ldots + k_N PSF\left(x, x_N\right)$$

In some texts the PSF is known as the 'impulse response function' (IRF).

Implicit in equation 2.24 is that an input impulse results in the same shaped output regardless of where it lands across the field of view: i.e. the system is linear and shift-invariant. For flat-panel digital detectors this is generally true, although the situation becomes more complicated for modalities involving reconstruction algorithms or heavy image processing, such as CT or MRI. However, many systems can be considered shift-invariant over a small area and the methodology described can still be successfully applied if a phantom can be designed such that a sufficiently large shift-invariant area is available for the measurement. Non shift-invariant systems may sometimes be corrected to be

so through measuring and compensating for the variation in response across the field of view (e.g. by uniformity corrections). Systems which are linear and shift-invariant are known as LSI and the position-independent PSF can be written

$$PSF\left(x, x_0\right) = PSF\left(x - x_0\right) \tag{2.25}$$

Digital images consist of pixels $p_{x,y}$ of finite size. In the 1-dimensional case, if the pixel size is Δx then an object $h\left(x\right)$ being imaged can be described by a series of rectangles of width $\triangle x$, centred on position i and height h_i, with the area of each rectangle thus $h_i \triangle x$. If $\triangle x$ is small relative to the width of the PSF then each rectangle can be represented by a delta function at that position, scaled by $h_i \triangle x$. Therefore, the output of a system with point-spread function $PSF\left(x, i\right)$ can be calculated by superposing the contributions from each of these delta functions. i.e.

$$S\left\{h\left(x\right)\right\} \approx \sum_{i=-\infty}^{\infty} h_i PSF\left(x, i\right) \triangle x \tag{2.26}$$

In the limit as the pixel size tends to zero, this becomes the superposition integral

$$S\left\{h\left(x\right)\right\} = \int_{-\infty}^{\infty} h\left(x'\right) PSF\left(x, x'\right) dx' \tag{2.27}$$

By virtue of equation 2.25, if the system is LSI this can be simplified to the convolution integral

$$S\left\{h\left(x\right)\right\} = \int_{-\infty}^{\infty} h\left(x'\right) PSF\left(x - x'\right) dx' \tag{2.28}$$

which may be represented as

$$S\left\{h\left(x\right)\right\} = h\left(x\right) \otimes PSF\left(x\right) \tag{2.29}$$

That is, the object signal is smeared out by the point-spread function.

In the above methodology, the performance of the system is described in the absence of noise. For real-world systems subject to stochastic noise, equations 2.28 and2.29 represent the expectation function of the system response. If deterministic 'fixed-pattern' noise is also present, this can be included by being added to the formalism independently. It can be thus sometimes be useful to

consider the response of an imaging system as:

$$S\left\{h\left(x\right)\right\} = \left\{h\left(x\right) + noise_{stoch}\left(x\right)\right\} \otimes PSF\left(x\right) + noise_{det}\left(x\right) \tag{2.30}$$

although care must be taken to avoid oversimplification because the different sources of noise may themselves also be functions of $h\left(x\right)$. A more detailed consideration of noise is presented below.

So far, the formation of the image has been considered entirely in the spatial domain. However, the convolution integral can be difficult to compute analytically and the ramifications of the smearing of differently sized structures by the PSF is unintuitive. It is therefore useful to consider signal transfer in the spatial frequency domain. An input sinusoidal signal of spatial frequency u may be expressed in terms of a complex exponential

$$h\left(x\right) = e^{i2\pi ux} = \cos\left(2\pi ux\right) + i\sin\left(2\pi ux\right) \tag{2.31}$$

According to equation 2.28 the resulting analogue image is thus

$$\begin{aligned} d\left(x\right) &= S\left\{h\left(x\right)\right\} \\ &= \int_{-\infty}^{\infty} PSF\left(x'\right) e^{i2\pi u\left(x-x'\right)} dx' \qquad (2.32) \\ &= e^{i2\pi ux} \int_{-\infty}^{\infty} PSF\left(x'\right) e^{-i2\pi ux'} dx' \qquad (2.33) \end{aligned}$$

Defining the Fourier Transform of a function $f\left(x\right)$ as

$$\mathcal{F}\left\{f\left(x\right)\right\} = F\left(u\right) = \int_{-\infty}^{\infty} f\left(x\right) e^{-i2\pi ux} dx \tag{2.34}$$

and the inverse Fourier Transform of $F\left(u\right)$ as

$$\mathcal{F}^{-1}\left\{F\left(u\right)\right\} = f\left(x\right) = \int_{-\infty}^{\infty} F\left(u\right) e^{i2\pi ux} du \tag{2.35}$$

then the Fourier Transform of the PSF is

$$\mathcal{F}\left\{PSF\left(x\right)\right\} = T\left(u\right) = \int_{-\infty}^{\infty} PSF\left(x\right) e^{-i2\pi ux} dx \tag{2.36}$$

and equation 2.33 becomes

$$\begin{aligned} d(x) &= e^{i2\pi ux}\mathcal{F}\{PSF(x)\} && (2.37)\\ &= T(u)\,e^{i2\pi ux} = T(u)\,h(x) && (2.38)\end{aligned}$$

In the general sense, for any specified input $h(x)$ with Fourier transform $H(u)$ the corresponding output will be

$$\begin{aligned} d(x) &= S\{h(x)\}\\ &= S\{\mathcal{F}^{-1}\{H(u)\} && (2.39)\\ &= \mathcal{F}^{-1}\{H(u)\} * PSF(x) && (2.40)\end{aligned}$$

Noting that convolution in the spatial domain corresponds with multiplication in the spatial frequency domain,

$$\begin{aligned} d(x) &= \mathcal{F}^{-1}\{H(u)\,\mathcal{F}\{PSF(x)\}\} && (2.41)\\ &= \mathcal{F}^{-1}\{H(u)\,T(u)\} && (2.42)\end{aligned}$$

Expressing $d(x)$ in terms of its Fourier Transform $D(u)$, yields the result

$$D(u) = H(u)\,T(u) \tag{2.43}$$

so that it is possible to characterise the change in a signal as it transfers through the system by multiplying each of its spatial-frequency components by the factor $T(u)$. It is thus straightforward to consider the transport of signals through the imaging system by working in the spatial-frequency domain. $T(u)$ is complex, so represents both linear scaling and a shifting of phase. However, if the PSF is real and even, which is generally the case for the imaging systems considered in this study, then $T(u)$ does not contain any phase information and $T(u) = |T(u)|$.

In the spatial domain, an oscillating signal is characterised by its amplitude and phase. The modulation M of the signal is a measure of its amplitude relative to its mean value, and is defined as

$$M = \frac{h_{max} - h_{min}}{h_{max} + h_{min}} \tag{2.44}$$

Consider the general case of equation 2.31, where the sinusoidal signal has

constant offset a and amplitude b:

$$h(x) = a + be^{i2\pi ux} \tag{2.45}$$

The modulation of this signal on input to the imaging system M_{in} is given by equation 2.44 as

$$M_{in} = \frac{b}{a} \tag{2.46}$$

After passage through the imaging system, the signal $h(x)$ is represented by an analogue image

$$\begin{aligned} d(x) &= S\{h(x)\} \\ &= S\left\{a + be^{i2\pi ux}\right\} \end{aligned} \tag{2.47}$$

In a linear system this becomes

$$d(x) = aS\{1\} + bS\left\{e^{i2\pi ux}\right\} \tag{2.48}$$

$$= aS\left\{e^{i2\pi ux|_{u=0}}\right\} + bS\left\{e^{i2\pi ux}\right\} \tag{2.49}$$

$$= aT(0) + bT(u)\, e^{i2\pi ux} \tag{2.50}$$

using the definition of $T(u)$ in equation 2.36. By equation 2.44, the output modulation is thus

$$M_{out} = \frac{b}{a}\frac{|T(u)|}{T(0)} = M_{in}\frac{|T(u)|}{T(0)} \tag{2.51}$$

with the ratio M_{out}/M_{in} being defined as the Modulation Transfer Function (MTF):

$$MTF(u) = \frac{M_{out}}{M_{in}} = \frac{|T(u)|}{T(0)} \tag{2.52}$$

which is the spatial-frequency domain equivalent of the contrast transfer factor T_c previously defined in equation 2.11. The MTF is always real, having lost all phase information because of the modulus operator in equation 2.52, and is defined as being unity at zero frequency. As described above, phase information is generally unimportant for the modalities being considered here, but if phase must be considered, then the equivalent of the MTF with phase information preserved is the Optical Transfer Function (OTF), given by:

$$OTF(u) = \frac{T(u)}{T(0)} \tag{2.53}$$

Therefore

$$MTF\left(u\right) = \left|OTF\left(u\right)\right| = \frac{\left|\mathcal{F}\left\{PSF\left(x\right)\right\}\right|}{\mathcal{F}\left\{PSF\left(x\right)\right\}|_{u=0}} \tag{2.54}$$

When working in 2-dimensions, the equivalents are

$$MTF\left(u,v\right) = \left|OTF\left(u,v\right)\right| = \frac{\left|\mathcal{F}\left\{PSF\left(x,y\right)\right\}\right|}{\mathcal{F}\left\{PSF\left(x,y\right)\right\}|_{u=0,v=0}} \tag{2.55}$$

In 2-d, the impulse response along one image axis may be studied by integrating the $PSF\left(x,y\right)$ along the other direction. This gives the line spread function (LSF), which represents the response of the system to an input which is an infinitesimally narrow straight line of infinite length. In each direction, the LSF is thus given by:

$$LSF\left(x\right) = \frac{\int_{-\infty}^{\infty} PSF\left(x,y\right) dy}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} PSF\left(x,y\right) dxdy} \tag{2.56}$$

and

$$LSF\left(y\right) = \frac{\int_{-\infty}^{\infty} PSF\left(x,y\right) dx}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} PSF\left(x,y\right) dxdy} \tag{2.57}$$

In the general case, the LSF normal to a vector $\mathbf{r}$ is given by

$$LSF = \frac{\int_{-\infty}^{\infty} PSF\left(\mathbf{r}\right) d\mathbf{r}}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} PSF\left(x,y\right) dxdy} \tag{2.58}$$

By these definitions, the modulus of the Fourier Transform of the LSF yields the corresponding MTF

$$MTF\left(u\right) = \left|\mathcal{F}\left\{LSF\left(x\right)\right\}\right| \tag{2.59}$$

and

$$MTF\left(v\right) = \left|\mathcal{F}\left\{LSF\left(y\right)\right\}\right| \tag{2.60}$$

These are equivalent to extracting the values along the u or v axes from the 2-d map described by 2.55, viz

$$MTF\left(u\right) = MTF\left(u,v\right)|_{v=0} \tag{2.61}$$

$$MTF\left(v\right) = MTF\left(u,v\right)|_{u=0} \tag{2.62}$$

Some important characteristics are shared between the spatial and spatial-frequency domains: if the PSF is real and even, then so is the MTF; if the PSF is

rotationally symmetric, then the MTF also shares this property.

Spatial domain approaches to describing noise were outlined in section 2.4, all of which involved measures of pixel variance either in the ensemble average of many instances of an image, or if the system is ergodic, within a region of interest. These are limited in that they are unable to characterise structure in the noise. Detailed analysis of noise structure can be performed by examining how signal fluctuations relate to one another as a function of time, or distance in an ergodic system.

Consider a noisy signal $a\left(x'\right)$ which fluctuates about its mean value at position x' by $\Delta a\left(x'\right)$ due to the noise. The autocovariance function $K\left(x', x'+x\right)$ then describes the covariance of $a\left(x'\right)$ with itself after offset by a distance x:

$$K_a\left(x', x'+x\right) = E\left\{\Delta a\left(x'\right)\Delta a^*\left(x'+x\right)\right\} \tag{2.63}$$

where $\Delta a^*\left(x'\right)$ is the complex conjugate of $\Delta a\left(x'\right)$. If the noise is *stationary*, such that its expectation value and variance do not vary as a function of position, then the autocovariance function is also independent of position across a uniform field of view

$$K_a\left(x', x'+x\right) = K_a\left(x\right) \tag{2.64}$$

Furthermore, in an ergodic system the autocovariance can be estimated by a spatial average over a limited field of view X, the *sample autocovariance*, $K_{a,X}\left(x\right)$

$$K_{a,X}\left(x\right) = \frac{1}{X}\int_X \Delta a\left(x'\right)\Delta a^*\left(x'+x\right)\mathrm{d}x' \tag{2.65}$$

in which the expectation function in equation 2.63 is replaced by a correlation integral.

As the field of view increases the sample autocovariance becomes a better approximation of the true autocovariance. In the limit as $X \to \infty$ the two become equal.

$$\begin{aligned} K_a\left(x\right) &= \lim_{X\to\infty} K_{a,X}\left(x\right) && (2.66)\\ &= \lim_{X\to\infty}\frac{1}{X}\int_X \Delta a\left(x'\right)\Delta a^*\left(x'+x\right)\mathrm{d}x' && (2.67)\end{aligned}$$

Although the autocovariance function provides a detailed characterisation of noise its calculation is computationally intensive and thus time-consuming. Furthermore, with most imaging modalities it is impractical to acquire suffi-

cient instances of a noise image to calculate $K_a(x)$ at a single position as in equation 2.64, or to achieve a uniform region of interest sufficiently large to produce a good estimate using the approach in equation 2.67. In addition, the autocovariance function itself provides results which are difficult to interpret in a meaningful context and it is not immediately straightforward to identify or separate out the contributions from different sources of noise. However, the situation is improved by working in the spatial frequency domain.

The Noise Power Spectrum (NPS) describes the noise variance resulting from different spatial frequency contributions and is the Fourier Transform of the autocovariance function

$$NPS(u) = \mathcal{F}\{K_a(x)\} \tag{2.68}$$

or more practically for an ergodic system

$$\begin{aligned} NPS(u) &= \mathcal{F}\left\{\lim_{X\to\infty} K_{a,X}(x)\right\} && (2.69)\\ &= \lim_{X\to\infty}\frac{1}{X}\mathcal{F}\left\{\int_X \Delta a(x')\,\Delta a^*(x'+x)\,\mathrm{d}x'\right\} && (2.70)\end{aligned}$$

which by the correlation theorem becomes

$$NPS(u) = \lim_{X\to\infty}\frac{1}{X}E\left\{\left|\mathcal{F}_X\{\Delta a(x)\}\right|^2\right\} \tag{2.71}$$

where $\mathcal{F}_X$ indicates that the Fourier Transform is performed over the limited range of size X.

Effectively, equation 2.71 is the Fourier Transform of a sufficiently large noise image, in which the background signal has been subtracted to yield an image containing values representative of the fluctuations due to noise only. It is relatively quick and straightforward to calculate due to readily available and efficient Fast Fourier Transform (FFT) algorithms. Uncertainty estimates in NPS calculations will be covered in section 2.6 but an accurate measurement tends to require averaging a number of individual results, either from multiple ROIs across a larger uniform field of view, or using multiple images acquired under the same operating conditions.

In any given imaging modality, noise is introduced at each stage of information transfer in the imaging chain. Noise from each source combines to give the total noise level in an image, as indicated by equation 2.30. This equation is a simplification, with the actual relationship between different sources of

noise being complex and often multiplicative. However, under low contrast conditions a linear transfer model is a good approximation and the equation holds. Detailed noise analysis therefore tends to be performed at an operating point where the noisy fluctuations in signal are relatively small in comparison with the signal itself. Under these conditions, obtaining a sufficiently large uniform field so that the changes in pixel value $\Delta a\left(x\right)$ are due to noise only is not always straightforward. Subtle trends in underlying signal may be present as a result of intrinsic characteristics of the modality, such as the heel effect in projection radiography, or sub-optimal experimental conditions, such as an imperfectly aligned detector. Unfortunately, these tend to be of a magnitude comparable with that of the noise fluctuations themselves, therefore biasing the NPS calculation. This effect can be somewhat mitigated by constructing a model of the signal trends in the region of interest and subtracting this prior to the noise analysis. Typical models include fitting a plane or 2D polynomial surface to the signal data in the region of interest. However, it is rarely possible to completely eradicate such low frequency trends and for this reason the lowest frequency results in an NPS calculation (i.e. those lying along the u and v axes) are often discarded.

Traditionally, noise analysis tends to focus on stochastic sources. However, it is useful to be able to separate out the contributions from stochastic and deterministic noise. Stochastic noise varies randomly across the field of view and between image instances. Although the exact fluctuations due to stochastic noise cannot be predicted the population can be described by distribution functions, the properties of which are themselves well defined. For example, stochastic noise from photon emission and detection is described by Poisson statistics, where the standard deviation of the distribution of photons is equal to the square root of its mean. Deterministic noise may be randomly distributed, but will appear in the same locations from one instance of the image to the next. Minimising the influence of deterministic noise may be achieved by identifying the pattern, such as the locations of defective pixels in an EPID or period and direction of image banding due to a readout synchronisation artefact, then applying a correction. Levels of stochastic and deterministic noise may themselves be a function of the underlying signal level, so consistency of operating conditions is important during the characterisation process. Under well-defined, low contrast conditions the noise contributions can be considered additive:

$$NPS_{total}\left(u\right) = NPS_{stoch}\left(u\right) - NPS_{det}\left(u\right) \tag{2.72}$$

$NPS_{total}(u)$ is determined using equation 2.71, estimating the expectation function by averaging as many individual realisations $NPS_i(u)$ as possible. Assessment of the contribution due to fixed pattern noise $NPS_{det}(u)$ can be performed by minimising stochastic noise through averaging many instances of the image, then calculating the power spectrum of this average image. The power spectrum of stochastic noise $NPS_{stoch}(u)$ is then the difference between the two

$$\begin{aligned} NPS_{stoch}(u) &= NPS_{total}(u) + NPS_{det}(u) && (2.73) \\ &= \frac{1}{N}\Sigma_{i=1}^{N} NPS_i(u) - NPS\left\{\frac{1}{M}\Sigma_{j=1}^{M} ROI_j(x)\right\} && (2.74) \end{aligned}$$

and is also sometimes referred to as the Wiener noise spectrum $W(u)$. This is generally the most useful noise descriptor because, in the absence of equipment failures, stochastic noise contributions tend to dominate in contemporary imaging modalities.

In two dimensions, equation 2.71 becomes

$$NPS(u,v) = \lim_{X,Y\to\infty} \frac{1}{XY} E\left\{\left|\mathcal{F}_X\left\{\Delta a(x,y)\right\}\right|^2\right\} \tag{2.75}$$

NPS is specified in units of variance per frequency bin, typically mm^2: i.e. $\left(1/mm^{-1}\right)\left(1/mm^{-1}\right) = mm^2$.

An important result, which arises from Parseval's Theorem[274], is that the area under the NPS curve is equal to the variance of the signal values in the $\Delta a(x,y)$ image. This is useful when validating numerical implementations of the NPS algorithm, as will be described in section B.18.

In the same way that equations 2.61 and 2.62 allow the MTF to be reduced into one dimensional functions containing contributions linked predominantly with the directions of the primary axes in frequency space, it is possible to reduce $NPS(u,v)$ into

$$NPS(u) = NPS(u,v)_{|v=0} \tag{2.76}$$

and

$$NPS(v) = NPS(u,v)_{|u=0} \tag{2.77}$$

using either slit based methods or by averaging bands of data around the primary axes.

When X-ray film was in widespread use, noise fluctuations could be measured

by scanning a uniformly exposed film using a scanning microdensitometer with a slit aperture orientated with its long axis perpendicular to the direction of travel. All light within the slit was integrated to give a single value at each measurement location. As the film passed through the microdensitometer the two-dimensional noise signal was thus collapsed into a one-dimensional function of position. The NPS corresponding to the direction of travel could then be calculated using equation 2.71.

For an infinitely long film moving in the x direction the 1D NPS given by this traditional *slit method* is[353]

$$NPS_S(u) = \int_{-\infty}^{\infty} S(u,v) \left|T(u,v)\right|^2 dv \tag{2.78}$$

where $S(u,v)$ is the Fourier Transform of the noise image $\Delta a(x,y)$ and $T(u,v)$ is the optical transfer function of the slit, representing the smearing effect of the finite integration area. If the slit is width w and length L then the point spread function $h(x,y)$ is given by

$$h(x,y) = \left(\frac{1}{w}\text{rect}(x,w)\right)\left(\frac{1}{L}\text{rect}(y,L)\right) \tag{2.79}$$

where $\text{rect}(a,b)$ is a rectangular aperture function which returns $1/b$ if $a < b$ and 0 otherwise. The optical transfer function is then given by

$$T(u,v) = \mathcal{F}\{h(x,y)\} = \text{sinc}(wu)\,\text{sinc}(Lv) \tag{2.80}$$

where the sinc function is defined as

$$\text{sinc}(x) = \frac{\sin \pi x}{\pi x} \tag{2.81}$$

Therefore, the NPS becomes

$$NPS_S(u) = \text{sinc}^2(wu) \int_{-\infty}^{\infty} S(u,v)\,\text{sinc}^2(Lv)\,dv \tag{2.82}$$

If the slit width is very narrow so that $w \to 0$ then $\text{sinc}(wu) \to 1$. Furthermore, if the slit length L is very long relative to its width, i.e. $L \gg w$ then $\text{sinc}^2(Lv)$ becomes sufficiently narrow that $S(u,v)$ is effectively integrated only over a short range about the v axis, where it is approximately constant. Under

these conditions,

$$NPS_S(u) = 1\int_{-\infty}^{\infty} S(u,v)\,\mathrm{sinc}^2(Lv)\,dv \tag{2.83}$$

$$\approx S(u,0)\int_{-\infty}^{\infty} \mathrm{sinc}^2(Lv)\,dv \tag{2.84}$$

$$= \frac{S(u,0)}{L} \tag{2.85}$$

and hence the 2D NPS is reduced to a 1D function. Care must be taken to ensure the slit length is long enough for the assumptions underpinning this formulism to be valid, otherwise low-frequency contributions to the NPS may be significantly underestimated[310]. However, in practice this may not be straightforward[353].

In digital (or non film-based) systems a similar methodology can be applied by recording a 2D noise image, summing the rows or columns to simulate scanning by a slit, then taking the Fourier Transform of the resultant 1D function. Although widely used in early studies[163, 187] this *synthetic slit* method has gradually given way to evaluations involving the whole 2D NPS, partly because 2D Fast Fourier Transform (FFT) algorithms are readily available and computing costs have become less significant, but also because this approach suppresses off-axis artefacts which are important to consider when performing a detailed noise characterisation of a system[76, 353].

If the 2D NPS has been calculated from a noise image using equation 2.75 then 1D representations in the directions of the primary axes can be calculated by averaging contributions from narrow bands around either axis. (The current IEC standard specifies ±7 pixels either side[146].) However, when undertaking this process it is important to calculate the frequency f of any off-axis point (u,v) as a quadratic sum $f = \sqrt{u^2+v^2}$ to ensure contributions at the same spatial frequency are being averaged.

Due to imperfect trend removal there may still be residual contributions along the zero-frequency axes themselves, so these are often omitted from the averaging bands. However, this has potential for obscuring structures lying on the axes themselves, so whether both axes are omitted or just the one orthogonal to the direction being considered depends upon the modality, the experimental conditions and confidence in the trend removal process. Necessarily this involves some trial and error so there is a role for standardisation across individual modalities and applications. However, it has been demonstrated that when the calculation algorithms are fully optimised the direct 2D technique

and synthetic slit method both yield equivalent results[353]. Because of this, and also the current predominance of 2D NPS techniques, the synthetic slit method is not considered further in this work.

Working with absolute noise levels can sometimes be problematic due to the dependence of stochastic noise on signal level. For example, if attempting to compare the noise levels of two portal imaging systems each unit may be limited to a small range of operating conditions which may be fundamentally different between the two – newer systems can acquire images with $\leq$1 MU, yet older systems require as many as 4 MU. Alternatively, the beam conditions, detector electronics and hence signal level will be very different between a radiographic X-ray examination and a megavoltage portal image, yet both are available for use in modern image-guided radiotherapy delivery systems and it may be desirable to compare the influence of noise between the two. There is also the somewhat unintuitive result that, due to Poisson statistics, an image formed from ionising radiation at a higher dose has a higher noise power than one at low dose – yet the perceived contrast is poorer!

Meaningful comparisons and interpretations can be assisted by dividing $NPS\,(u,v)$ with the square of the mean of the signal $S\,(s,y)$ in the region (or image instances) being considered, yielding the normalised noise power spectrum (NNPS) in units of mm^2.

$$NNPS\,(u,v) = \frac{NPS\,(u,v)}{|S\,(x,y)|^2} \tag{2.86}$$

In the spatial-frequency domain analysis above, MTF therefore provides a detailed assessment of spatial resolution and NPS (or NNPS) a measure of system noise. During an optimisation exercise the effect on either of these quantities can be examined whilst varying modality operating parameters. However, it is attractive to combine the two into a single metric which provides an overall indicator of imaging performance across the spatial frequency spectrum.

The Noise Equivalent Quanta (NEQ) is given by

$$NEQ\,(u,v) = \frac{T_c^2 MTF\,(u,v)^2}{NPS\,(u,v)} \tag{2.87}$$

where T_c is the large-area transfer factor given by equation 2.11. If the system is linearised such that the output signal is zero when the input signal is zero (i.e. there is no systematic offset in the system transfer curve) then equations 2.86

and 2.86 can be combined to give

$$NEQ\,(u,v) = \frac{MTF\,(u,v)^2}{NNPS\,(u,v)} \tag{2.88}$$

thus providing a straightforward means of combining resolution and noise characteristics into a single metric.

NEQ is the number of quanta falling on a perfect detector which would yield the same SNR as that observed in the real system under consideration, at the specified operating point. Effectively, it is a measure of the number of quanta the image is 'worth'. NEQ is an important quantity because it is representative of the total information content in the image and therefore has a role in human observer models, as described in section 2.7 below.

Whereas NEQ is an absolute measure of image quality, an absolute indicator of overall system performance can be achieved by expressing NEQ as a function of 'exposure' in terms of the number of quanta Q involved in the image acquisition process. Dividing NEQ by Q yields the Detective Quantum Efficiency (DQE)

$$DQE\,(u,v,Q) = \frac{NEQ\,(u,v)}{Q} = \frac{T_c^2 MTF\,(u,v)^2}{Q.NPS\,(u,v)} \tag{2.89}$$

which effectively expresses overall performance (in terms of spatial resolution and noise power) as a function of imaging dose. If the object being imaged can be modelled such that the signal to noise ratio $SNR_{in}\,(u,v)$ input to an ideal, photon counting detector can be deduced, then DQE can also be expressed as

$$DQE\,(u,v,Q) = \frac{SNR_{out}^2\,(u,v,Q)}{SNR_{in}^2\,(u,v)} \tag{2.90}$$

where $SNR_{out}\,(u,v,Q)$ is the signal to noise ratio observed in the final image at the specified operating point. DQE therefore enables technical measurements of imaging system performance to be linked back to the fundamental signal detection theory underpinning the clinical task. However, one caveat is that because the MTF and NPS may themselves vary with dose, measurement conditions and the spectral components of the incident radiation, it cannot strictly be assumed that $DQE\,(u,v,Q)$ is linear with Q over anything other than a narrow range of values. Therefore, when describing equipment performance in terms of DQE the experimental conditions must always be carefully specified, especially when drawing comparisons between different instances of the same modality.

Although it can be instructive to examine 3D graphs of $NEQ(u,v)$ and $DQE(u,v)$ it is usually easier to apply the results if they are reduced to predominantly u and v components, as described above for MTF and NPS. If the MTF and NPS results are binned to the same spatial frequency intervals then the curves can be directly combined using equation 2.89 above.

In the diagnostic imaging community NEQ and DQE are generally accepted as 'gold standard' measures of image quality and are widely used to perform objective comparisons of examples of the same modality being applied to a particular task. For example, a great deal of work has been undertaken on comparative assessments of imaging plates for computed radiography, or imaging plates for digital radiography. However, their role in radiotherapy optimisation is less certain. On its own DQE can be used to optimise the performance of a single modality, but it cannot directly be applied to the optimisation of multiple modalities, all with fundamentally different operating principles and exposure parameters, against each other for a complex multi-step task. It may help inform purchasing decisions - for example, when a range of virtual simulation systems are available, all producing DRRs, and a relative performance assessment is desirable. However, radiotherapy equipment tends to be bought as part of an overall package, with image quality for one particular part of the solution being only one consideration. (DRR quality depends upon CT acquisition parameters, so the CT scanner and virtual simulation package must be assessed together.) It is also unclear how DQE should be applied to modalities not involving ionising radiation, such as ultrasound and MRI, because the number of quanta Q in equation 2.89 cannot easily be specified.

NEQ is perhaps more useful in radiotherapy optimisation, allowing overall performance to be assessed at different equipment operating points. For example, the NEQ can be measured for 'standard' (low dose) and 'high quality' (high dose) EPIs so that the clinical benefit from the additional patient dose can be assessed. However, as will be discussed in section 2.7, there is as yet no robust observer model for specific radiotherapy tasks, let alone the overall optimisation process, with the result being that NEQ curves are difficult to interpret and apply to a clinical context. Whereas diagnostic imaging involves a constant balancing of dose, spatial resolution and noise performance, in radiotherapy imaging the important factors are much more task dependent. For example, when optimising EPIs used for patient set-up verification, if the DRRs being used to match against are only accurate to 3 mm, then there is little advantage in improving the spatial resolution characteristics of the EPID

much beyond this, with more benefit being gained from improving the noise characteristics. Furthermore, a factor of 4 increase in imaging dose between 0.5 MU and 2 MU is still a tiny fraction (~1%) of the overall dose from the treatment, so there is considerably more flexibility in dose than there is for diagnostic imaging.

For these reasons, until adequate observer models have been developed, it is suggested that MTF and NPS are more directly applicable to radiotherapy image quality evaluation and optimisation. Both can intuitively be interpreted in the appropriate clinical context and weighted in the optmisation process according to their perceived importance. Although the optimisation process itself will still be subjective, the metrics used to inform the process are robust, detailed and can be directly linked to elements of the clinical task. In the examples provided in the later chapters of this thesis, MTF and NPS are therefore considered in preference to NEQ and DQE, although the IQWorks package can be used to calculate any of these as appropriate for a particular application.

2.6. Image Quality Metrics for Digital Systems

A *digital image* represents the mapping of a continuous function $d(x, y)$ onto an array of discrete picture elements (*pixels*) $p_{x,y}$, with centre-to-centre spacings Δx and Δy in each matrix direction and overall dimensions $N\Delta x$ and $M\Delta y$. For some modalities (such as digital radiology (DR) or EPI) there is a direct correspondence between the elements of the pixel matrix and those of a two-dimensional detector array employed to acquire the image data. For others, such as CT, MRI, DRRs and PET, complex reconstruction algorithms may be required to construct an image from the measured detector signals. However, regardless of the details of the image formation process, similar theoretical concepts apply if the system can be considered linear and shift invariant. For this reason the term 'pixel' is used interchangeably with the concept of a detector element in the discussion below, although it is acknowledged that the two are not necessarily equivalent.

Mapping between the spatial and spatial-frequency domains is performed using the Discrete Fourier Transform $\mathcal{DF}$ and its inverse $\mathcal{DF}^{-1}$. For simplicity, but without loss of generality, in 1D these are defined as

$$D_{u,v} = \mathcal{DF}\{d_{n,m}\} = \sum_{n=0}^{N-1}\sum_{m=0}^{M-1} d_{n,m} e^{-i2\pi(nu/N+mv/M)} \tag{2.91}$$

$$d_{n,m} = \mathcal{DF}^{-1}\{D_{u,v}\} = \frac{1}{NM}\sum_{u=0}^{N-1}\sum_{v=0}^{M-1} D_{u,v}e^{i2\pi(nu/N+mv/M)} \tag{2.92}$$

where $n, u = 0, 1, 2, \ldots, N-1$ and $m, v = 0, 1, 2, \ldots, M-1$. These are similar to their continuous, integral-based counterparts in equations 2.34 and 2.35. However, an important difference is the way in which the locations of the input and output values are linked. The dimensions N and M of the input and output arrays are preserved, and if the samples in the spatial domain are at $n\Delta x$ and $m\Delta y$ in each direction, then the output values will be at spatial frequencies given by $\frac{u}{N\Delta x}$ and $\frac{v}{M\Delta y}$.

Ideally, a digital representation of an image would exactly match the analogue case

$$p_{n\Delta x, m\Delta y} = d(x,y)_{|x=n\Delta x, y=n\Delta y} \tag{2.93}$$

but this is rarely the case both due to the finite pixel size and ramifications of the sampling process. It therefore cannot be assumed that $D_m = D(u)$ at locations $u \in \frac{m}{N\Delta X}$.

As described in section 2.5, transfer through a system of both signal and noise amplitude is described by the optical transfer function (OTF). Transfer of an arbitrary signal through a digital detector may therefore be modelled as

$$P_d(u,v) = D(u,v)\,OTF_d(u,v) \tag{2.94}$$

where $D(u,v)$ and $P_d(u,v)$ are the spectral decompositions of the input signal and final digital image respectively. $OTF_d(u,v)$ is the digital optical transfer function of the system and can be resolved into a number of components.

Prior to sampling, the input signal will have been blurred by geometrical factors (e.g. the focal spot size in an x-ray system or linear accelerator), the physics of the imaging system (e.g. an x-ray photon being converted into a cascade of light photons in a phosphor, then subsequent scatter of these) and the aperture function of the imaging device (e.g. the finite area of the detector element, scanning laser spot size, or ray line width in a DRR). All of these combine to give the *presampling* optical transfer function, OTF_{pre}:

$$OTF_{pre} = OTF_{geom} \cdot OTF_{phys} \cdot OTF_{aper} \tag{2.95}$$

The aperture function is important because each detector element integrates signal contributions from across its whole sensitive area, rather than a sharp

point at its centre. If the detector elements are assumed to be perfect rectangles, then the point spread function $h(x,y)$ is described by

$$h(x,y) = \left(\frac{1}{\Delta x}\text{rect}(x,\Delta x)\right)\left(\frac{1}{\Delta y}\text{rect}(y,\Delta y)\right) \tag{2.96}$$

and OTF_{aper} is given by

$$OTF_{aper}(u,v) = \mathcal{F}\{h(x,y)\} = \text{sinc}(u\Delta x)\,\text{sinc}(v\Delta y) \tag{2.97}$$

Sampling by a digital detector is equivalent to multiplying the continuous presampled signal by the comb function $\mathbf{III}(x,y;\Delta x,\Delta y)$, which describes an infinite array of delta functions separated by Δx and Δy in the two directions of the matrix. Because the Fourier Transform of a comb function is another comb function $\mathbf{III}(u,v;1/\Delta x,1/\Delta y)$ with delta functions at reciprocal spacings in frequency space, the optical transfer function for the digital system becomes

$$OTF_d(u,v) = OTF_{pre}(u,v) \otimes \mathbf{III}\left(u,v;\frac{1}{\Delta x},\frac{1}{\Delta y}\right) \tag{2.98}$$

which effectively causes the OTF_{pre} to be replicated throughout frequency space on grid points separated by Δx^{-1} and Δy^{-1}. In the same way, the sampled noise image is given by

$$NPS_d(u,v) = NPS_{pre}(u,v) \otimes \mathbf{III}\left(u,v;\frac{1}{\Delta x},\frac{1}{\Delta y}\right) \tag{2.99}$$

According to the *Sampling Theorem*, if a continuous function is bandlimited so that it contains no frequency components above some critical frequency f_c then it can be completely described by regular sampling at an interval $\Delta = 2f_c$. However, if the function contains frequency components greater than f_c these will be wrapped around so that their contributions overlap with those of frequencies less than f_c. Because the overlapping contributions are indistinguishable from the true contributions at the lower frequencies, this effect is known as *aliasing*. The critical frequency f_c is known as the *Nyquist frequency* and a signal is said to be *undersampled* if the sampling rate is insufficient to avoid aliasing.

Applying this to equations 2.98 and 2.99, the Nyquist frequencies in the two matrix directions are $u_c = 1/(2\Delta x)$ and $v_c = 1/(2\Delta y)$, and any frequency components in the input signal above this will be subject to aliasing. Figure 2.4 demonstrates the ramifications of aliasing on equation 2.98, where the one dimensional case is considered for simplicity. In figure 2.4a OTF_{pre} has

no components above u_c and there is no overlap between its replications in frequency space. The measured samples at frequencies u_1 and u_2 are therefore representative of the presampled analogue signal. However, when components greater than u_c are present, as in 2.4b, there is overlap between replications, with overlapping contributions summing to give an effective OTF_d magnitude indicated by the dotted line. In this scenario u_1 is faithfully transferred, but u_2 is significantly overestimated. Aliasing clearly influences the direct measurement of MTF and NPS and care must therefore be taken to minimise its influence when attempting to draw objective comparisons between different systems.

A further complication arises because convolution by the comb function in equations 2.98 and 2.99 stops the system being shift invariant if it is undersampled, unless the shift is a multiple of Δx or Δy along one of the primary axes. This does not matter for NPS assessment because the frequency components of stochastic noise are uncorrelated so that the natural and aliased variances will add together in the same way regardless of phase, and the ensemble averaging in equation 2.75 incorporates contributions from all phases. However, it is particularly problematic for MTF measurements where a deterministic signal is being interpreted and the frequency components are correlated and phase dependent.

One solution is to directly measure the presampled MTF by designing the experiment in such a way that the signal used to calculated the MTF is oversampled[73, 100, 332]. This can be achieved by orientating a line or edge at a slight angle to one of the axes of the detector matrix so that a step of one pixel along this axis results in a much smaller change in distance from the line or edge. For example, if an edge or line is angled close to the y axis then equation 2.98 is reduced to

$$OTF'_{pre}(u) = OTF_{pre}(u) \otimes \mathbf{III}\left(u; \frac{1}{\Delta x'}\right) \tag{2.100}$$

where the supersampling rate $1/\Delta x'$ is essentially so high that $OTF'_{pre}(u) \approx OTF_{pre}(u)$, with negligible aliasing. This technique therefore successfully extracts the 1D MTF_{pre} corresponding to the matrix axis perpendicular to the edge. It is widely employed today and its numerical implementation is described further in section B.17.3.

However, there is an argument that MTF_{pre} does not adequately reflect true detector performance because aliasing (and phase dependency) will still be present in real medical images acquired when the system is in clinical use.

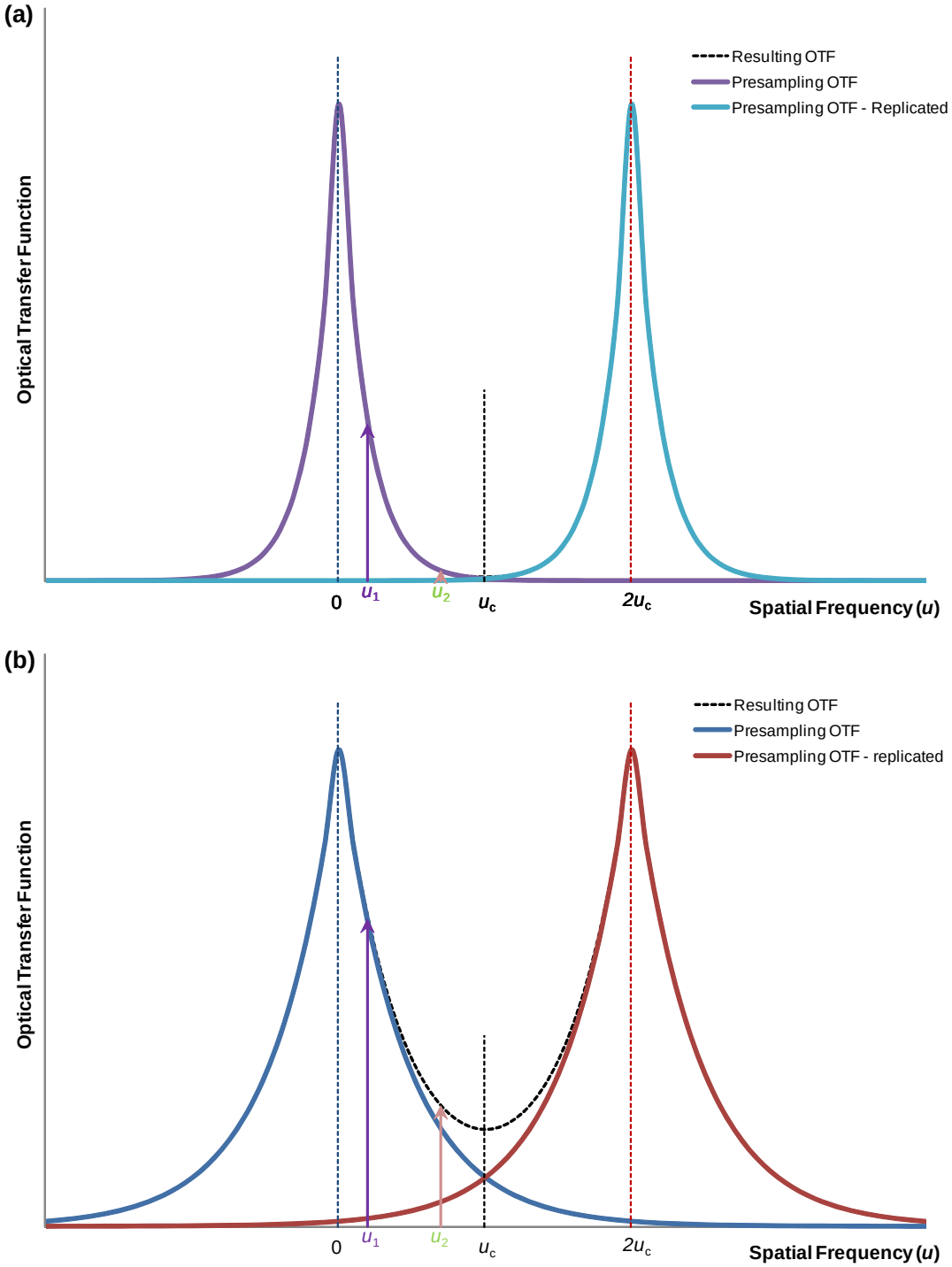


Figure 2.4. – Illustration of the effect of aliasing on the pre-sampled Optical Transfer Function (OTF_{pre}) of a signal sampled at a rate of $2u_c$. (a) The frequency spectrum of OTF_{pre} contains no frequencies above u_c, so there is no overlap when OTF_{pre} is replicated in frequency space at an interval of $2u_c$. (b) Frequencies above u_c are present, so that the sampling rate of $2u_c$ is insufficient to prevent overlap of the replicated OTF_{pre}s and the frequencies above u_c are folded back to be below u_c. In this instance, the magnitude of frequency component u_2 will be overestimated due to aliasing.

A more comprehensive analysis of MTF can be performed by first calculating MTF_{pre} then substituting this back into equation 2.98 (using the relation $MTF_{pre}(u,v) = \left|OTF_{pre}(u,v)\right|$) and examining the minimum and maximum MTF_d which might be observed depending upon the phase relationship of the comb function with the pixel matrix. In addition, a global measure of MTF can be obtained by averaging MTF_d over all phases[74]. This *Expectation MTF* or EMTF is shift invariant generally agrees with MTF_{pre} at frequencies close to zero, but tends to overestimate it closer to the Nyquist frequency, sometimes by as much as 30%. Furthermore, it is difficult to calculate, requiring numerical integration of an already known or measured MTF_{pre}, and does not describe the transfer of a specific frequency sinusoid at any specific location so its interpretation is difficult. For these reasons it is not widely used in image quality evaluation today and will not be considered in detail in this work. However, its phase-averaging nature means that it may closely resemble the MTF measured using repeating bar patterns[80, 124, 381], a technique which is widely employed and will be discussed in section B.17.2.

Given the discussion above, it is clear that calculations of NEQ and DQE using equations 2.88 and 2.89 will give very different results depending upon exactly how the MTF and NPS are determined, and particularly whether the signal is undersampled. In terms of stochastic noise aliasing, as long as the input frequency spectrum is representative of that found in patient imaging (generally broad spectrum noise) then the directly measured sampled NPS, including aliasing, will be representative of that encountered in the clinical situation. Regarding MTF assessment, a shift invariant system is required to facilitate a stable measurement, so phase dependency and aliasing should be avoided. Therefore, there is general consensus in the literature[76, 146, 308] that it is most appropriate to use the presampled MTF and the sampled NPS in these calculations, after rebinning these functions to be at the same frequency interval.

An assumption in equations 2.98–2.100 is that an infinite array of delta functions is used to sample the image. However, in reality the sampling window will always be a finite region of interest, with the maximum possible dimensions being those of the image itself. Sampling within a finite region is equivalent to multiplying the input signal by a spatial windowing function, $w(x,y)$. Multiplying by a window function in the spatial domain corresponds to convolution by the Fourier Transform of this function $W(u,v)$ in frequency space, so that

the equations become

$$OTF_d(u,v) = OTF_{pre}(u,v) \otimes \mathbf{III}\left(u,v;\frac{1}{\Delta x},\frac{1}{\Delta y}\right) \otimes W(u,v) \tag{2.101}$$

$$NPS_d(u,v) = NPS_{pre}(u,v) \otimes \mathbf{III}\left(u,v;\frac{1}{\Delta x},\frac{1}{\Delta y}\right) \otimes W(u,v) \tag{2.102}$$

$$OTF'_{pre}(u) = OTF_{pre}(u) \otimes \mathbf{III}\left(u;\frac{1}{\Delta x'}\right) \otimes W(u) \tag{2.103}$$

If a simple rectangular region is being considered, of dimensions X and Y, then the window function is simply unity everywhere within the ROI and zero elsewhere

$$w_{rect}(x,y) = \left(\frac{1}{X}\text{rect}(x,X)\right)\left(\frac{1}{Y}\text{rect}(y,Y)\right) \tag{2.104}$$

which has the Fourier Transform of a 2D sinc function

$$W(u,v) = \text{sinc}(uX)\,\text{sinc}(vY) \tag{2.105}$$

The side lobes and central peak of the sinc function result in any particular frequency bin containing aliased contributions from nearby bins, as illustrated in Figure 2.5. This *leakage* between bins can significantly bias MTF and NPS measurements, particularly at low spatial frequencies[73]. Simple rectangular sampling windows cause particularly wide-band leakage because their sharp edges contain a broad spectrum of frequency components. This can be reduced by applying artificial windowing functions which gradually reduce the signal value towards zero at the edges of the region of interest. These act as low-pass filters which suppress high frequency (aliased) contributions whilst allowing lower frequency (non-aliased) contributions to pass. When windowing a region used for MTF estimation the convention is for the window function to be normalised to unity at its centre, although in practice this does not matter because the MTF calculation itself is self-normalising. However, in NPS calculations it is important that the windowing function is normalised so that its root-mean square is unity, in order to preserve the noise power.

In the context of data windowing, the rectangular window (i.e. effectively applying no additional window function) is known as the Dirichlet window. A range of smoother data windows are in common use[274], and those implemented in IQWorks are indicated below. These are described as a 1D function of pixel index j across a sampling region of width N pixels. If applied in 2D they are rotationally symmetric about the centre of the ROI.

The Ramp (or Bartlett) window

$$w_{bartlett,j} = 1 - \left(\frac{j - N/2}{N/2} \right) \tag{2.106}$$

The Cosine window

$$w_{cosine,j} = \cos\left(\frac{\pi j}{N-1} - \frac{\pi}{2} \right) = \sin\left(\frac{\pi j}{N-1} \right) \tag{2.107}$$

The Hamming window

$$w_{Hamming,j} = 0.54 - 0.46\cos\left(\frac{2\pi j}{N-1} \right) \tag{2.108}$$

The Hann window

$$w_{Hann,j} = 0.5 - 0.5\cos\left(\frac{2\pi j}{N-1} \right) \tag{2.109}$$

The Welch Window

$$w_{Welch,j} = 1 - \left(\frac{j - N/2}{N/2} \right)^2 \tag{2.110}$$

The Gauss Window

$$w_{Gauss,j} = e^{-\frac{1}{2}\left(\frac{j-(N-1)/2}{\sigma(N-1)/2} \right)} \tag{2.111}$$

where $\sigma \leq 0.5$. (In IQWorks $\sigma = 0.5$.)

Each window function is plotted in figure 2.6a for MTF calculations (normalised to unity at zero distance) and 2.6b for NPS calculations (normalised so that the RMS is unity). A comparison of the leakage resulting from each window is presented in figure 2.5. When choosing a window function there is generally a trade-off between the sharpness of the central peak and the magnitude of the side lobes: the sharper the central peak, the better the avoidance of close-range aliasing, whilst the smaller the side lobes the less the long-range aliasing. It is evident in figure 2.5 that, in comparison with the rectangular window, all other window functions have considerably smaller side lobes, albeit at the expense of broader central peaks. All of the above functions have been applied to MTF and NPS calculations in the literature, with the choice of which window is used for any particular application seeming rather arbitrary (see, for example [73, 76, 116, 214, 217, 264, 308]). Although the nature of the low-pass filter undoubtedly influences MTF and NPS results it is suggested that in practice the

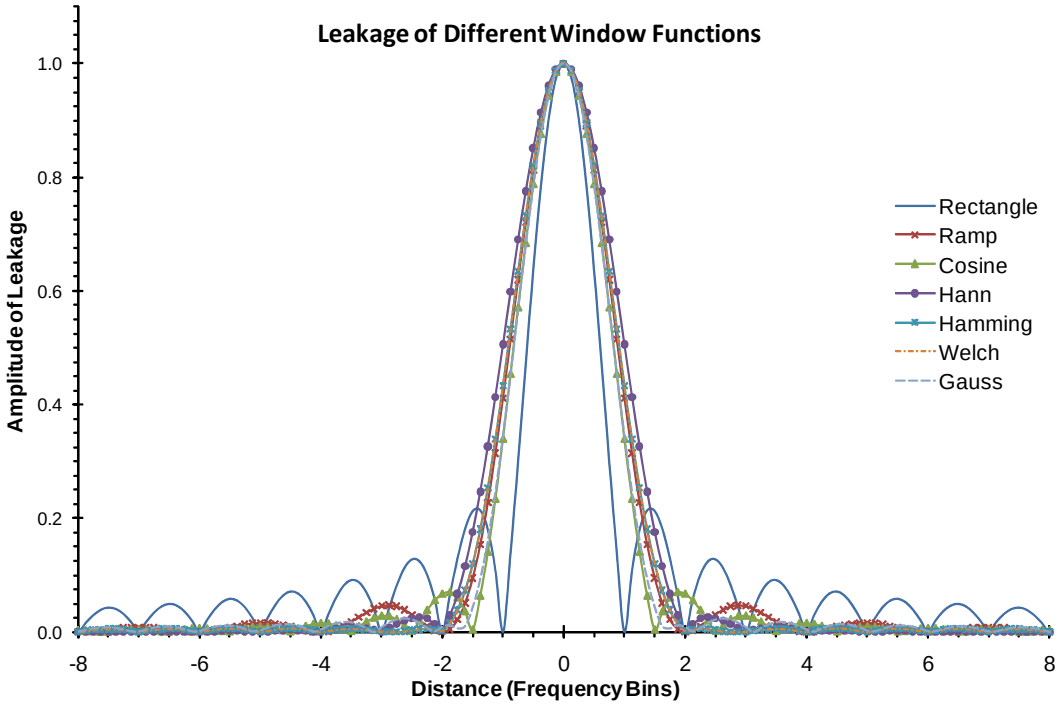


Figure 2.5. – A comparison of the leakage characteristics of the different window functions implemented in IQWorks.

precise choice of window is unimportant, as long as it is not rectangular[274]. However, recent investigations have indicated that consistent application of the window function is important when validating numerical MTF and NPS algorithms[179]. All the above window functions are available in IQWorks to enable users to apply whichever window has been cited in literature relevant to their application.

Clearly, the practical realisation of a window function depends upon the size of the ROI, and hence so too does its leakage characteristics. As the sampling region increases, the finer spacing of the frequency bins in the Discrete Fourier Transform results in a compression of the leakage function towards the limit of a delta function at zero frequency. Therefore, the larger the sampling region, the smaller the leakage due to the sampling window. It has been suggested that it can be more efficient to just increase the sampling region until the window leakage is negligible rather than applying any additional window function[73]. Indeed this is the approach followed by the IEC standard on DQE measurements[146]. However, this is not viable for modalities with relatively small pixel matrices, such as CT or EPI, so the availability of additional window functions is still important.

An interesting result when calculating the NPS of a digital image is that the

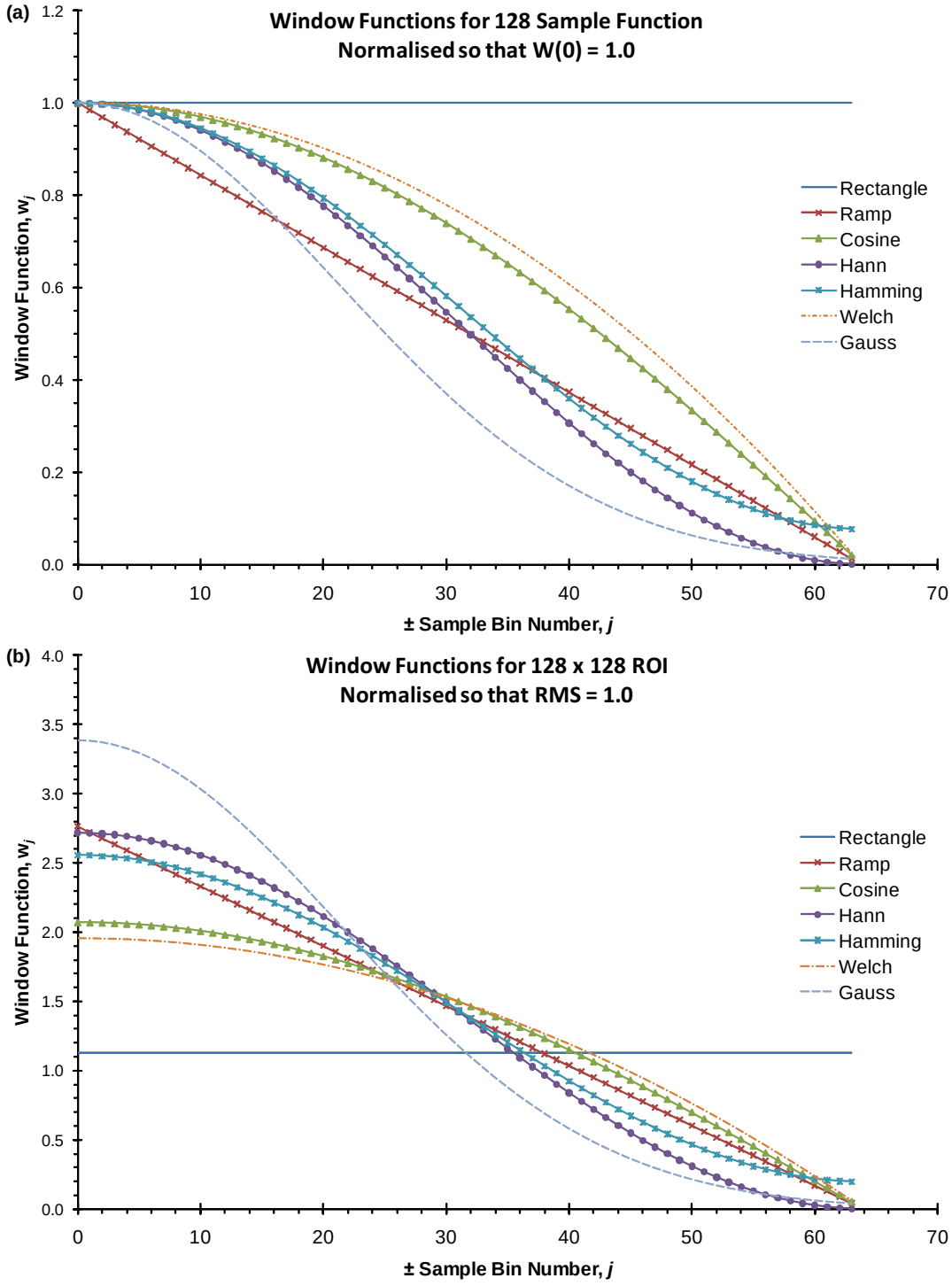


Figure 2.6.
Data window functions implemented in IQWorks, applied to a region of width 128 samples. a) Normalised to unity at the centre of the region of interest, for MTF measurements. b) Normalised so that the RMS value is unity, for NPS calculations.

sum of NPS_d is equal to the variance of the noise image used in the calculation

$$\sum_{u=0}^{N-1}\sum_{v=0}^{M-1} NPS_d(u,v) = \sigma^2_{image} \tag{2.112}$$

Furthermore, because there is an inherent uncertainty in any measured pixel value of an image due to stochastic noise there will also be an uncertainty in the noise power calculated in any particular frequency bin. Because the standard deviation of the noise power estimated in any bin is proportional to the actual power in that bin, the coefficient of variance (CoV) is independent of frequency. It is estimated by[73, 76]

$$CoV \approx \frac{1}{\sqrt{N_{bins}}} \tag{2.113}$$

where N_{bins} is the number of independent bins used in the noise power estimate. Therefore, as the number of averaged realisations of the noise in equation 2.75 increases, the CoV decreases accordingly. If overlapping ROIs are employed then the estimated CoV is increased in proportion to the relative areas used. For half-overlapping ROIs this is by a factor of $\sqrt{2}$ and for quarter-overlapping ROIs a factor of 2. These results are useful in validating digital NPS calculation algorithms and will be discussed in this context in section B.18.

2.7. Observer Models

Optimisation of an imaging study, or indeed the whole radiotherapy imaging process, requires the ability to convert technical measures of image quality into performance metrics which are clinically relevant. This involves understanding how a human observer perceives and interprets real images to aid the completion of specific tasks. Various models have been developed which take technically measurable parameters as input and calculate the effective signal to noise ratio seen by the observer. Some of these are relatively sophisticated, based on modelling of the human visual system and large observer trials, but central to all is the clear definition of the task being performed.

The simplest task is the identification of a completely specified region of known signal against a completely specified background: i.e. a signal region of known area and magnitude against a background of known area and magnitude, or else two regions of signal, again of known area and magnitude, lying on the same background. This is referred to as the Signal Known Exactly / Background Known Exactly (SKE/BKE) task and in a clinical context is analog-

ous to determining whether a tumour or anatomical structure of exactly known dimensions and signal properties can be discriminated in an image. It is clearly a gross oversimplification of a real clinical task, but is useful to explore because it forms the basis of more sophisticated models.

An early model of the SKE/BKE task in radiographic imaging was developed by Rose[44, 301–303] and calculated the detectability of a uniform region of area A against a uniform background of the same area. If the mean numbers of quanta per unit area falling on the region and background are $\overline{q_o}$ and $\overline{q_b}$ respectively then the contrast, given by equation 2.8, is

$$C = \frac{\overline{q_o} - \overline{q_b}}{\overline{q_b}} \tag{2.114}$$

Defining the 'Rose signal' as the change in the number of quanta due to the presence of the object, integrated over the area of the object, then this is equal to

$$\Delta S_{Rose} = (\overline{q_o} - \overline{q_b})\, A \tag{2.115}$$

If the background noise is both uncorrelated and Poisson distributed then the standard deviation of the noise σ_b is equal to the square root of the number of quanta in the background region

$$\sigma_b = \sqrt{A\overline{q_b}} \tag{2.116}$$

Substituting equations 2.115 and 2.116 into equation 2.5 yields

$$\Delta SNR_{Rose} = \frac{\Delta S_{Rose}}{\sigma_b} = \frac{A\,(\overline{q_o} - \overline{q_b})}{\sqrt{A\overline{q_b}}} = C\sqrt{A\overline{q_b}} \tag{2.117}$$

which is the classical definition of the Rose model[62]. As indicated in section 2.2, studies have shown that ΔSNR_{Rose} should generally be above around 3 – 5 in order for an object to be discriminated, depending upon how strictly an observer is scoring an image against what they are expecting to see[157, 218, 246, 301, 315].

It is clear from equation 2.117 that detectability depends upon both contrast magnitude and object size. This premise forms the basis of contrast-detail experiments using test objects comprising patches of different sizes and calibrated contrasts. In a detailed analysis the observer identifies details which are only just detectable against the background noise and constructs a curve of limiting object size against limiting contrast. Such contrast-detail curves may contribute

to the more involved ROC analyses described in section 2.2[141, 189, 221]. There is evidence that the principles of the Rose model are generally sound[50, 62] and its results describe the observed quantum limiting nature of noise in modalities involving ionising radiation. However, it is limited by the assumption that all noise in the system is Poisson distributed and uncorrelated, which is generally not the case. Correlations are introduced when stochastic noise is transferred through different stages of the imaging system (for example, when x-ray photons are converted to light photons in a phosphor, which are then detected by an amorphous-silicon semiconductor) and sources of noise which are non-Poisson distributed, such as electronic or film-grain noise may also be present. Although refinements of the Rose model are possible, for example to account for imperfections in the human visual system[118], it still suffers from the limitations of the other spatial domain metrics described above.

Working in the spatial-frequency domain, the simplest model of a hypothesis based decision maker is that of the *Bayesian* observer[141, 332], who calculates the likelihood of different outcomes and chooses the most probable one. This represents the *ideal* observer, who is able to extract all possible information from an acquired image and utilise fully all available prior information when performing the detection task. Effectively, the ideal observer perceives the original object faithfully except as degraded by the fundamental physics of the acquisition hardware. In terms of the spatial-frequency domain metrics identified in section 2.5, the SNR of the Bayesian observer is given by

$$SNR^2_{ideal} = T_c^2 \int\int \frac{MTF(u,v)^2 \Delta S\,(u,v)^2}{NPS\,(u,v)} \mathrm{d}u\mathrm{d}v \tag{2.118}$$

where $\Delta S\,(u,v)$ is the magnitude of the spatial-frequency domain representation (i.e. the Fourier Transform) of the difference in signal between the object and background, or between the two objects being considered. This model has the advantage over the simple Rose model in that it describes performance across the whole spatial-frequency spectrum and can be extended to more than just the SKE/BKE task. However, specifically for the SKE/BKE task, the Bayesian observer works by randomising structured noise present before seeking an object with the expected signal difference. This *prewhitening* of the noise has led to the Bayesian observer being known as a Pre-Whitening Matched Filter (PWMF).

Real human observers are able, to some extent, to 'see through' stochastic noise but find deterministic sources more difficult to deal with. Another model

intended to better match this performance by only being able to compensate for stochastic noise is the Non Pre-Whitening Matched Filter (NPWMF), described by

$$SNR^2_{NPWMF} = \frac{\left| \int\int \Delta S\left(u,v\right)^2 MTF\left(u,v\right)^2 \mathrm{d}u\mathrm{d}v \right|^2}{\int\int \Delta S\left(u,v\right)^2 MTF\left(u,v\right)^2 \left[MTF\left(u,v\right)^2 NPS_{obj}\left(u,v\right) + NPS\left(u,v\right) \right] \mathrm{d}u\mathrm{d}v} \tag{2.119}$$

where $NPS_{obj}\left(u,v\right)$ is the power spectrum of the structured noise in the object itself.

Both the PWMF and NPWMF models are able to access all information in the image, limited by the presence of noise. Although their formulisms are relatively straightforward, they have the disadvantage that they rely upon accurate prior knowledge of the object signal to be detected so that $\Delta S\left(u,v\right)$ can be calculated. In practice, this may be extremely difficult or impossible preventing application of these models to all but the simplest scenarios.

An extension of the PWMF is the Hotelling observer[14, 15, 130], which is more flexible in that it does not require knowledge of the object to be detected, does not place stringent requirements on the imaging conditions and can be applied directly to more than just the SKE/BKE task. Furthermore, it is thought to be closer to a human observer in that it is restricted to performing only linear operations on the image data. The figure of merit used by the Hotelling observer is

$$SNR^2_{Hotelling} = \int\int \frac{\Delta\widetilde{\widetilde{S}}\left(u,v\right)^2 MTF\left(u,v\right)^2}{MTF\left(u,v\right)^2 NPS_{obj}\left(u,v\right) + NPS\left(u,v\right)} \mathrm{d}u\mathrm{d}v \tag{2.120}$$

where $\Delta\widetilde{\widetilde{S}}\left(u,v\right)$ is the difference in the spatial-frequency domain spectra of the object and background (or the two objects), averaged over all possible outcomes. Both $\Delta\widetilde{\widetilde{S}}\left(u,v\right)$ and $NPS_{obj}\left(u,v\right)$ can be determined by performing iterative operations on recorded images, although the process may be computationally intensive. However, the Hotelling observer has the advantage of being capable of being applied to a range of clinically representative tasks where detailed prior information may be unavailable[141].

More advanced models have also been developed to take into account of noise and the MTF of the human visual system itself[245], or the effect of ambient lighting. For example, the ideal observer degraded by the resolving power of the human visual system is given by

$$SNR^2_{ideal+vis} = T_c^2 \frac{\left| \int\int \left| \Delta S\left(u,v\right) MTF\left(u,v\right) O\left(u,v\right) \right|^2 \mathrm{d}u\mathrm{d}v \right|^2}{\int\int \left| \Delta S\left(u,v\right) MTF\left(u,v\right)^2 O\left(u,v\right)^2 \right|^2 NEQ\left(u,v\right)^{-1} \mathrm{d}u\mathrm{d}v} \quad (2.121)$$

where $O\left(u,v\right)$ is the MTF of the observer's visual system[13, 218].

Regardless of the degree of sophistication, common to all these models is the way in which they utilise the MTF and NPS concepts to model the imaging system, visual system of the observer and the object being imaged. Work is ongoing to include refinements such as simulating how the human eye seeks out details in a pattern[189], modelling the radiation source or to include perturbations in the system caused by the presence of the object itself (e.g. scattered radiation or due to its biasing of the energy spectrum of the x-ray beam)[194]. It is also thought that the presence of normal anatomical structures in images complicates any detection task by introducing a non-uniform background[189]. Also, certain lesions and other abnormalities may also tend to be associated with particular anatomical structures, so that a human observer may be biased by their personal clinical experience or knowledge of the patient's medical history.

As radiotherapy imaging modalities become increasingly sophisticated, and multimodality image fusion more readily available, the complexity and time required for image interpretation also increases, along with the volume of imaging information to consider[262, 287]. Furthermore, with highly developed control systems it is possible to vary dose delivery in near real-time based on continuously acquired images, thus achieving better target conformation and providing the opportunity for dose escalation and improved tumour control. For all these reasons there is a drive in radiotherapy to automate image interpretation at different stages of the radiotherapy process. Success has been achieved in developing methods for automatic segmentation of the target volume and organs at risk[6, 32, 254, 256, 323], and imaging techniques for tracking and compensating for tumour position during treatment are also available[3, 10, 48, 320, 326, 342]. However, much of this research has been based on the availability of prior information (such as anatomical atlases or expert systems) or involves modelling the task using methods other than those described above. To date, there has been little work to develop the spatial-frequency domain observer models well used in diagnostic imaging to radiotherapy tasks, or indeed to simulating the entire radiotherapy process, yet the same basic principles apply. This is definitely an opportunity for the future and will be invaluable in developing

overall optimisation strategies.

Due to the lack of observer models for radiotherapy imaging tasks this study will involve calculating individual metrics such as MTF and NPS. These can then be fed into observer models which are developed in the future.

2.8. Geometrical Factors

Geometrical accuracy is paramount in radiotherapy, as discussed in section 1.5, yet is not traditionally included in discussions of image quality. Almost any radiotherapy task involving images relies on the assumption that the geometry being visualised is correct, or in the case of image fusion, that the geometry of at least one set of reference images can be relied upon. Because of its importance it is argued in this work that geometrical accuracy should be considered a fundamental image quality factor for radiotherapy images.

An image is geometrically accurate if dimensions of and distance vectors between structures are faithfully reproduced in 2D (or 3D if dealing with a volumetric image or stack of 2D images. There are two factors:

Pixel/Voxel Size: The separation between the centres of adjacent elements in a 2D / 3D image matrix. This may correspond to the sampling aperture size of the detector, but in digital imaging this cannot be assumed because the image is often processed before storage or presentation. As discussed above, pixel size is related to but not the same as spatial resolution. In most imaging modalities the pixel size is homogeneous across the whole extent, in both matrix directions, of a 2D image. When dealing with a 3D image, the pixel size tends to be uniform in each plane of the stack, with the depth dimension potentially being coarser, but this depends on the modality.

Geometric Linearity: Distances between points in the image are proportional to those measured on the original object. i.e There is no inherent distortion in the image. If pixel/voxel size is used as a scaling factor then distances should agree between the original object and the image.

Pixel/voxel size must be correct and the image must be geometrically linear for the distance vectors to be faithfully reproduced and the image to be geometrically accurate. For simplicity, consider a single 2D image plane. If the pixel size is wrong in one direction, then both the angle and magnitude of distance vectors will be incorrect. However, if it is wrong by the same proportion in both

directions (or, as is relatively common, the pixel size is arbitrarily set to unity in both directions) then the angles of distance vectors will be correct, but not the magnitude. Alternatively, even if the pixel size is correct, the image may not be geometrically linear because structures may be distorted (by an affine or non rigid-body transformation) when the image is formed. If linearity is not preserved, then neither are the angles and magnitudes of the distance vectors.

Although not specifically accounted for in the metrics described above, they are still affected by geometrical accuracy. The formulism for MTF assumes the system is shift-invariant, which itself relies upon geometric linearity. MTF is also inherently linked to pixel size, because this is used to calculate the intervals between points on the frequency axis. NPS does not use dimensional information per se, but the number of quanta contributing to a pixel depends upon the pixel area, and hence noise power is sensitive to pixel size. Furthermore, when calculating NNPS the convention is to normalise by pixel area, so again there is an effect. The result is that, although Fourier-based image quality metrics are influenced by geometrical accuracy it can be difficult to deduce to what extent or how any particular instance of a metric has been affected just by considering the metric itself in isolation. i.e. Fourier metrics are sensitive to geometrical inaccuracy but non-specific. Instead, dedicated methods are required to consider geometrical integrity.

If accurately aligned, most contemporary X-ray-based imaging modalities using solid-state flat panel detectors tend not to be susceptible to geometrical errors. Experience has shown that pixel/voxel size and geometric linearity will either be correct, or unambiguously wrong, with the most common causes being failures of mechanical alignment systems (e.g. when 'levelled' a detector is not actually orthogonal to the beam CAX so that the geometrical projection varies across the surface of the detector) or communication breakdown between the detection electronics and the image processing / storage system (such that pixel size information is lost or corrupted). For these modalities, it is very straightforward to assess geometrical accuracy by measuring distances between identifiable key points. Both the angle and magnitude of the distance vector can be verified by making measurements of distances crossing both axes of the pixel matrix. A number of measurements is required, both short and long range, to ensure geometrical distortion is not present.

If a modality is inherently susceptible to geometric distortion, such as MRI or Nuclear Medicine, then a more detailed analysis is required, taking images of test objects containing points on a defined grid, then calculating the displace-

ment vectors of each point. The vector map can then be compared against a baseline or used in a model to correct distortion and restore linearity. Unfortunately, the nature of geometrical distortion can be modality, acquisition setting and test object specific, so comprehensive assessment and correction schemes must be tailored to the application. Because the simpler approach outlined above is applicable to the majority of radiotherapy modalities, and is a sufficient QC check for all modalities, only that is considered in this work.

An important consideration in radiotherapy imaging is that all distances in images acquired by projection-like modalities (such as simulator images, DRRs, EPIs) are presented as if they were projected back along the beamline to the plane of the isocentre. This is usually performed automatically by the modality and is checked as part of the QA programme. However, it is a fundamentally different situation to that in diagnostic imaging, where dimensions are usually specified at the detector plane. The result is that, when calculating metrics such as MTF or NNPS which are sensitive to pixel size it is necessary to ensure the same geometrical corrections are being applied to all systems. To avoid ambiguity, all such calculations presented in this work are performed in the isocentre plane, unless specified otherwise.

2.9. Practical Measurement of Image Quality Metrics

Ultimately, the absolute test of imaging system performance is to compare decisions and measurements made by real human operators against the results of pathological or histological investigations. However, except in a very few cases[101], this is impractical. Furthermore, a detailed assessment of fundamental performance requires the measurement of the objective image quality metrics described above under controlled conditions.

Measurement of a particular metric is performed by acquiring images of a suitable test object or *phantom* constructed to be conducive to the measurement task and ideally representative of the clinical situation being modelled. It is important to note that whilst many phantoms are somewhat abstract in appearance and may not appear to a human observer to resemble anything found in the clinical environment, the key is that fundamental (and therefore deliberately abstract) parameters are measured under conditions similar to those found in real patients, such as dose to the phantom and detector, x-ray beam spectrum, range of contrasts in the image, etc.

Each of the metrics described earlier in this chapter can be assessed using relatively simple, well-defined phantom components. Regardless of the modality being considered the same basic principles apply when determining specific performance metrics, so the components used to perform these assessments tend to be applicable across modality. For example, uniformity and NPS evaluations are performed on simple uniform regions of phantoms (or on flood-field exposures)[31, 75, 146]; MTF can be calculated using an impulse point object[33, 170], a line[63], an edge[167, 307] or a bar pattern[80, 124, 381]; geometrical linearity is evaluated by measuring the distance between identifiable structures across the field of view.

In this study the aspiration was to employ the most appropriate assessment method to each modality being investigated, introducing techniques and adapting phantoms previously developed for other modalities if required. Examples are presented in chapters 4 to 5 where phantoms not previously associated with particular radiotherapy imaging modalities have been introduced because they provide a more detailed or efficient assessment than previously.

When comparing instances of the same modality it is important to maintain consistency in phantom construction, the geometrical set-up of the experiment, exposure conditions and other fundamental factors which might affect image quality. IEC standards exist prescribing how DQE assessment of diagnostic detectors should be undertaken[146] and describing standard beam conditions to use when reporting results of other experiments[148]. Whilst useful as a starting point, and essential when quantitative performance metrics are being agreed as part of a contract, such standardisation may not be appropriate for clinical optimisation. For example, the beam arrangement in the IEC DQE standard involves the use of secondary collimators, which is uncommon in diagnostic imaging, and requires a very specific beam filtration which can only be achieved through a time-consuming, iterative process and the use of very expensive high purity aluminium. This means that the set-up is impractical to achieve in a routine clinical setting, and may not be representative of real clinical conditions. Furthermore, the standard experimental conditions cannot be directly mapped to the measurement of DQE of megavoltage portal imaging systems due to the significantly different beam spectra involved and it being extremely difficult at these penetrating energies to manufacture phantoms which are truly radio-opaque, a requirement of the standard.

It is therefore suggested that for optimisation of clinical protocols experimental conditions are chosen which are straightforward to replicate yet mimic

the clinical scenario as closely as possible. Furthermore, for radiotherapy imaging modalities it is important to consider how the measurements made at one stage of the patient pathway can be compared with those performed elsewhere. For example, it would be advantageous to ensure the same scanning protocol was used both during assessment of image quality for target volume delineation and when considering DRRs, and also ideally the same phantom, and that these results are compatible with those obtained for kV or MV verification imaging. For quality control tests a pragmatic approach should be adopted where objective metrics are calculated under 'reasonably' representative conditions. Whilst not suitable for fundamental performance evaluation, the results produced would still be robust and objective metrics valid for routine constancy testing.

Quantities such as MTF and NPS provide a considerable amount of information which enables a detailed picture of the performance of an imaging system to be constructed. However, for the purposes of quick QC tests, or for gross comparisons between systems, they can sometimes be too complex to enable meaningful conclusions to be efficiently drawn. For this reason, it is useful to be able to reduce the functional quantities into two or three numerical values which provide summary indices of image quality. For example, when calculating MTF the full curve tends only to be examined when performing a detailed performance evaluation. Instead, it is standard practice to quote the 'critical' spatial frequencies at which the curve falls to a few specified fractions of unity (i.e. fractions of $MTF\,(0)$), such as f_{50}, f_{10} and f_2, corresponding to the frequencies at which the MTF curve is 50%, 10% and 2%. The key is to balance the information lost by data reduction with the practical usefulness of the quality index. Further examples of reduced indices of image quality are discussed when the numerical algorithms are presented in appendix B.

One final consideration is the 'quality' of the data contained within the images being analysed to measure image quality. For example, if an edge image for an MTF measurement is particularly noisy, or a supposedly uniform field for a NPS assessment contains significant non-uniformities (e.g. due to saturation artefacts), then the standard numerical algorithms may not cope well with the non-ideal data. In particular, the differentiation of an edge image to determine the LSF for MTF calculation is robust only when the edge signal varies monotonically with distance, so that even small amounts of noise can make this operation challenging[214]. Additional image processing is therefore often required to 'condition' image data to be suitable for input into a standard algorithm. This might include curve fitting, local averaging or other data

reduction methods. Care must be taken to ensure data conditioning does not inappropriately influence the final results although, as discussed above, it is suggested that a pragmatic approach be adopted that is in line with the intended application. Data conditioning methods implemented in IQWorks are discussed in chapter B.

2.10. Conclusion

In the same way that a 'good quality' image is one which is suitable for the clinical task, a good image quality metric is also task dependent. Although the radiotherapy imaging task has not yet been modelled, there is strong evidence that the calculation of standardised metrics underpinned by fundamental imaging science will benefit any optimisation exercise.

Large-area metrics provide a gross, global assessment of performance, whereas spatial-frequency domain metrics such as MTF and NNPS yield considerably more information regarding how a signal is transferred through the system, and are thus more indicative of the strengths and limitations of particular systems. Whilst large-area metrics are quick and easy to calculate, giving straightforward quantitative results, spatial-frequency domain metrics tend to be vector quantities which need to somehow be reduced to single numerical values if they are to be applicable to QC checks. Spatial-domain metrics are clearly more appropriate when performing fundamental evaluations of absolute imaging performance. Overall, a consideration in both domains is required to fully characterise a system.

In radiotherapy, related modalities may be so fundamentally different that it is necessary to adopt a pragmatic approach in which useful metrics can be practically measured on different systems which are as equivalent and clinically relevant as possible. For example, when comparing the performance of a DRR or simulator image (an early step in the patient pathway) with a megavoltage EPI (a dependent, later step in the pathway) it is suggested that for optimisation purposes it is more prudent to consider clinically representative radiation beams and detector doses than perhaps those recommended in the literature for absolute performance characterisation. Indeed, a possible hindrance to the adoption of quantitative assessment techniques in both diagnostic and radiotherapy imaging is the practically difficult experimental techniques often described in the literature which simply cannot be realised in basic QC tests. As demonstrated in chapters 3–5, it is still possible to harness the benefits of

quantitative, fundamental and standardised image quality metrics by taking a pragmatic approach appropriate to the task concerned. Furthermore, it is suggested that these metrics have a place across the whole spectrum of radiotherapy tasks, including technique and technology development, routine QA and optimisation.

IQWorks as a software framework to facilitate the automated and objective evaluation of image quality is discussed in Appendix A. Details of the numerical algorithms utilised to implement the metrics described in this chapter then follow in Appendix B and a summary list of all metrics calculated by IQWorks is included in Appendix C. If more than one equation has been presented for any particular metric in the discussion above, then the user can interactively choose which equation to use, depending upon their requirements. A key design goal of IQWorks was to keep the framework as flexible as possible, giving the physicist the freedom to choose whether to exactly mimic a particular experimental scenario and analysis methodology from the literature, or instead implement an acceptable compromise. Crucially, it is straightforward in IQWorks to change from one calculation algorithm to another so that the influence of different implementations can be quickly evaluated, and the robustness of particular techniques understood.

Chapter 3.

Modality 1: Electronic Portal Imaging

3.1. Overview

This chapter applies the image analysis framework developed over the previous chapters, and practically implemented via the IQWorks package described in appendices A and B, to the performance evaluation of electronic portal imaging devices (EPIDs).

Following a description of the linacs and EPIDs employed in this study, two commercial phantoms for the assessment of EPID image quality are considered: the QC-3V phantom[281] (formerly of Masthead Imaging Corp, British Columbia, Canada, now Standard Imaging Inc, Wisconsin, USA) and the PTW EPID phantom[275] (PTW, Freiburg, Germany). Both phantoms are in widespread use in the radiotherapy community[123, 204, 269] and are marketed with bespoke software to quantitatively analyse image acquired using them: PIPS Pro[372, 381] with the QC-3V phantom, and epidSoft[275, 370] with the PTW phantom. Analysis trees were developed in IQWorks to mimic the operation of the commercial packages, both to explore the implementation of their analysis routines but also as part of an intercomparison based validation of all three packages. If IQWorks can be shown to generate results similar to those of commercial packages then it can potentially be used as a direct replacement, enabling the flexibility and extensibility of IQWorks to be applied to more advanced assessments and experiments in the future whilst maintaining compatibility with historical consistency measurements. All metrics calculated by the commercial packages were examined using IQWorks, including modulation transfer function (MTF), geometric linearity, contrast, contrast to noise ratio (CNR) and local signal linearity. Interesting anomalies in the behaviour of both

commercial packages are revealed by this exercise.

A new EPI performance phantom is also introduced which was developed as part of this work, the 'QEPI1' phantom. Building upon techniques utilised for performance evaluation in diagnostic imaging[76, 308], this offers the same analysis opportunities as the commercial phantoms but contains additional features to facilitate the assessment of fundamental equipment performance across the field of view. Development and testing of this phantom for routine use in the Oxford Cancer Centre imaging quality assurance (QA) programme is described.

If images of all three portal imaging phantoms are acquired under similar experimental conditions, then analysed using equivalent numerical methods, then the performance metrics calculated should be comparable. However, intrinsic differences in the implementations of the commercial packages prevent this being the case. These are discussed in this chapter and IQWorks analysis trees are introduced which aim to fulfil the intent of the original phantom designers yet are consistent in algorithmic approach across all the phantoms.

A detailed comparison of the two currently available models of Varian EPID is then performed using the QEPI1 phantom. These devices – the aS500-II and aS1000 detectors – are described in section 3.2, and the intention of the comparison was to determine whether the enhanced features of the more expensive detector, the aS1000, provide tangible benefits for the clinical set-up verification task. Both MTF and normalised noise power spectrum (NNPS) are considered in the study. Results of a more general intercomparison of different generations of EPID, in terms of the frequency at which the MTF drops to 50% (f_{50}) and CNR, are also discussed.

Following the commissioning of the new Oxford Cancer Centre, the QEPI1 phantom has formed the basis of the EPID QA programme in which it is used both for quick daily QC checks and for more detailed troubleshooting and optimisation. Performance trends are described and the routine application of the phantom is discussed.

3.2. Linacs and Detectors Considered in this Work

IQWorks and the QEPI1 phantom were developed by the author over a number of years whilst working in the Edinburgh and Oxford Cancer Centres. Over this period measurements were performed using eleven Varian linear accelerators (Varian Medical Systems, Palo Alto, California, USA), details of which are

presented in table 3.1. Two different models of electronic portal imaging device (EPID) were used to to acquire images: in Edinburgh, Varian's older aS500 unit[350], which is no longer available for sale, and in Oxford the newer aS1000 device[347]. Summary technical specifications of the three models are included in table 3.2, along with those of a third model – the aS500-II – which is sold as a less expensive, more limited functionality version of the aS1000 and has been included for comparative purposes. From the citations listed in the table it is clear that whilst a great deal of work has been published on the aS500 detector, less has been undertaken with the more recent models.

Essentially, all three models are based around the same flat panel detector, with the differences between them lying in the digitisation and processing electronics[107, 229, 322, 343, 347, 350]. The sensitive layer of the detector is an amorphous silicon (aSi) panel divided into an active matrix of 1024 columns and 768 rows over an area of 40×30 cm^2. Each cell of the matrix corresponds to one detector element and is a square of side 0.39 mm that incorporates a thin-film transistor (TFT) to allow charge integrated in the cell to be read out. Optically coupled to the aSi layer is a ~0.5 mm thick (134 mg/cm^2) high atomic number gadolinium oxysulphide scintillating phosphor (Lanex Fast Back screen – Eastman Kodak, Rochester, New York, USA) which converts incoming photons and electrons to light, greatly improving the sensitivity of the detector[248]. Additional build-up is provided by a 1 mm thick copper plate that converts X-ray photons into Compton electrons and which brings the total equivalent depth of the aSi sensitive layer to 8 mm of water[92]. This boosts the image signal at the expense of a slight degradation of spatial resolution due to blurring caused by increased Compton scatter in the metal plate[182, 183].

Charge is collected by each matrix element when the panel is exposed to ionising radiation. Following exposure, the detector matrix is read out line by line, with the charges in all elements in any particular line being addressed and read out simultaneously. After amplification, the analogue charge values are converted to a digital signal by a 14 bit analogue to digital converter (ADC) then this signal is transmitted along a serial link to a processing unit which generates the raw image. This is passed to the clinical imaging application which performs any additional image processing before displaying the image to the user. Together, the detector panel, readout electronics, ADC and subsequent processing unit constitute the Image Acquisition System (IAS), which is at version 2 in the aS500 detector and version 3 in the aS500-II and aS1000.

Linac Name	Centre	Linac Model	Photon Energies (MV)	Imaging Dose Rate (MU/min)	Treatment Dose Rate (MU/min)	EPID Model	OBI?
LA1	ECC	600CD	6	100	500	aS500	No
LA2	ECC	600C	6	80	400	aS500	No
LA3	ECC	6EX	6	100	500	aS500	No
LA4	ECC	21EX	8, 15	100	500	aS500	No
LA5	ECC	2100CD	8, 15	100	500	aS500	No
V1	OCC	21iX	6, 15	100	600	aS1000	Yes
V2	OCC	21iX	6, 15	100	600	aS1000	Yes
V3	OCC	21iX	6, 15	100	600	aS1000	No
V4	OCC	21iX	6, 15	100	600	aS1000	No
V5	OCC	21iX	6, 15	100	600	aS1000	No
V6	OCC	21iX	6, 15	100	600	aS1000	No

Table 3.1. – Linear accelerators and electronic portal imaging devices (EPIDs) considered in this study. In the 'Centre' column, ECC and OCC denote Edinburgh and Oxford Cancer Centres respectively. OBI refers to Varian's kilovoltage 'On-Board Imaging' option.

EPID Model	Acquisition System		Matrix Size	Pixel Size (mm)	Frames per Image		Dose per Image (MU)		Refs
	Version	Mode			SQ	HQ	SQ	HQ	
aS500	IAS2	Sync	512 × 384	0.78	4	10	3-4	7-9	[24, 108, 229, 230, 232, 322, 350]
aS500-II	IAS3	Rad Shot	512 × 384	0.78	2	4	1	2	[85, 86, 343, 347]
aS1000	IAS3	Rad Shot	1024 × 768	0.39	2	4	1	2	[227, 347]

Table 3.2. – Technical specifications of the electronic portal imaging devices (EPIDs) considered in this study. All have an active detector panel size of 40 cm × 30 cm. SQ and HQ refer to standard and high quality modes respectively.

Whereas the IAS2 relies on a dedicated processing computer with its own operating system and local storage to process the image data, the IAS3 utilises control and frame grabber boards in the linac control computer, which is a conventional IBM compatible PC running Microsoft Windows (Microsoft Corp, Redmond, Virginia, USA). This results in a more flexible and integrated system in which communication between the different components is simplified. Greater lower-level access to the acquisition control systems allows more powerful image processing applications to be developed than was possible with the previous IAS version. For example, calibration images are stored directly on the main linac control computer and can be accessed by Windows software directly in version 3, whereas these were stored remotely on the bespoke processing computer in version 2[347, 350].

A key functional difference between the two generations of IAS is that the newer model has considerably faster readout and processing electronics[260]. These allow every detector element to be addressed directly in the IAS3, whereas in the IAS2 it was necessary for them to be binned 2×2 in order for them to be read out in a timely manner. There is therefore potential for a significant boost in spatial resolution with the newer system due to the finer effective pixel size.

Another important difference is how charge is read out from the panel. In the IAS2, a so-called 'Sync' mode is employed in which rows of the panel are read out in between linac beam pulses. Essentially, immediately following each pulse of radiation a specified number of rows is read out from the detector. This cycle is repeated until the whole panel has been read out (constituting one 'frame'), then again until the required number of frames has been acquired, with the average of all acquired frames being returned to the processing system as the image. In order for a similar dose to be delivered to each group of rows it is important that the dose-rate is constant throughout the entire acquisition process and this is achieved by the IAS taking control of the modulation of beam pulses, synchronising their delivery against an internal clock. Nevertheless, the system is still susceptible to instabilities when the beam is first switched on and these are overcome by including 'reset frames' at the start of an acquisition which are read out as normal then discarded before the real imaging frames are acquired, resulting in images which do not suffer from beam stabilisation artefacts. By default, the system is configured to have two reset frames, although some researchers have investigated whether this can be reduced to one[232, 350]. Apart from the dose penalty due to the 'wasted' reset frames, limitations of the Sync mode are that horizontal line artefacts may be introduced by glitches in beam

synchronisation, especially if the underlying timing of the pulse modulation has drifted, and that the number of rows which can be read out between pulses depends upon the pulse repetition frequency. i.e. A panel can be read out more quickly and easily for a 100 MU/min beam than a 600 MU/min one because the time gap between pulses is longer. However, in practice this limitation is not generally an issue because 100 MU/min is routinely utilised for imaging of set-up verification fields, as described below.

With the IAS3 a new 'Rad Shot' mode was introduced that makes use of a feature of the linac that inhibits exposure on a particular control signal. This functionality was originally intended to facilitate gated treatments, allowing the treatment beam to remain 'turned on' but suppressing electron gun pulses so that no radiation is delivered. Because the linear accelerator remains energised X-rays can be switched on and off rapidly without the dose-rate stabilisation issues encountered when energising the whole system from cold. When using Rad Shot mode the panel is exposed for a defined period of time, then the beam is held off whilst the panel is read out in its entirety[347]. This removes the need to acquire reset frames whilst radiation is being delivered and completely eliminates timing or dose-rate stabilisation artefacts because the entire panel is exposed to the same variations in beam conditions. In practice, the exposure time is varied by the acquisition software depending upon the specified dose-rate, such that a nominal constant dose is delivered per frame, regardless of the dose-rate set. In Oxford the nominal Rad Shot dose per frame was set at the default 0.5 cGy / frame, in line with that utilised by other researchers[227, 343].

Together, the new Rad Shot mode and the ability of the faster IAS3 electronics to read out every cell in the detector matrix, have the potential to considerably improve image quality, at a lower patient dose[42, 343, 347]. However, as indicated in table 3.2, whereas the Rad Shot mode is available as standard in both the aS500-II and aS1000 EPIDs, the aS500-II is limited to 2×2 pixel binning so has an intrinsically lower spatial resolution. Although the two models are sold as different devices, experience has shown that the difference between the two is a licence file provided by Varian which enables 'full resolution' readout. By removing this file the aS1000 units in Oxford act as aS500-II units, and this is how the relative performance of the two modes is evaluated in section 3.7.

When specifying the dose required to acquire an image it is important to follow a methodology compatible with other researchers and which can be mapped to the patient imaging process. As discussed in section 1.2, radiotherapy linacs are calibrated so that one chamber monitor unit (1 MU) corresponds

to a dose of 1 cGy measured in a full-scatter water phantom on the beam central axis at the depth of dose maximum for a 10×10 cm^2 field at a source-surface distance (SSD) of 100 cm. This means that for an EPID SSD of 100 cm the dose at the sensitive aSi layer from 1 MU of radiation is of the order of 1 cGy, although this varies with beam energy (because the effective buildup of 8 mm water is generally less than the depth of dose maximum) and field size. This is similar to the situation encountered within the patient during a treatment: the geometric centre of the patient is usually at isocentre, and the dose at isocentre for any given beam will also be approximately 1 cGy/MU, although this is heavily dependent on the size of the patient, the tissues in the irradiated volume, field size and beam energy. Precisely quantifying the dose to isocentre, the dose distribution within the patient and the dose at the portal imager is relatively complex, and it is attractive to employ a simpler model when focusing on the impact of detector technologies or configuration changes on image quality. Therefore, patient doses from portal imaging are often specified simply in terms of MU, with it being assumed that 1 MU is approximately 1 cGy to the patient at isocentre. This is the convention adopted by Varian when specifying a nominal dose per frame for the IAS3 'Rad Shot' mode, as described above, and leads to imaging doses being specified in terms of MU in table 3.2. It is understood that in detailed optimisation exercises it is important to consider patient doses in a more detailed fashion using one of the approaches discussed in section 1.8.

In this study, all test images were acquired under standard geometrical conditions: a detector SSD of 140 cm (the Varian standard for imaging calibrations and measurements[347, 350]) and a field size sufficient to cover the surface of the detector (28.5×21.4 cm^2). Taking the inverse square law into account, this results in a dose to the detector of approximately 0.5 cGy / MU, with the exact dose depending upon the beam energy and nature of the phantom present in the beam.

All models of Varian EPID can acquire images in Standard Quality (SQ) and High Quality (HQ) modes with the difference between them being the number of frames averaged, and hence the duration and overall dose of the exposure. Although the number of frames is user-customisable the Varian defaults were used in Oxford and Edinburgh and these are indicated in table 3.2, along with the number of monitor units required to acquire an image. It can be observed that the dose per image is somewhat variable for the IAS2 based units because the frame readout time is dose-rate dependent, as described above, so is susceptible to differences in tuning of the IAS based dose-rate

control systems[107, 347, 350].

Eight out of the eleven linacs considered in this work could deliver two nominal energies of X-ray beam, with the other three able to deliver only a single 6 MV beam. All linacs were also capable of delivering a number of dose-rates, ranging from 80 MU/min to 600 MU/min. However, in practice a single dose-rate was always chosen to deliver all but very specialised treatment beams: on all but one of the linacs in Edinburgh this was 500 MU/min, the second highest available dose-rate, with the exception being the older generation linac LA2, where the highest available dose-rate of 400 MU/min was utilised. In Oxford, the highest available dose-rate of 600 MU/min was always employed for treatment beams. With Varian EPIDs any combination of photon beam energy and dose-rate can be utilised to acquire an image, enabling treatment beams to be imaged without modification. However, especially for more complex treatments, the projections provided by treatment fields may not be ideal for the purposes of assessing set-up errors. Generally, anterior-posterior / posterior-anterior and lateral projections are the most desirable when assessing set-up errors for two reasons: they are those which radiographers and doctors learn to interpret during their clinical training, and they inherently allow errors to be resolved into longitudinal, lateral and vertical vectors which can be directly corrected by couch motions. Therefore, if no treatment fields are available at appropriate gantry angles additional imaging fields may be added to provide these projections. Indeed, for workflow and consistency purposes, additional anterior-posterior and right lateral imaging fields are added routinely to most radiotherapy treatments in Oxford.

In both Edinburgh and Oxford all dedicated imaging fields are delivered at the lowest available energy and dose-rate. Reducing the beam energy increases the likelihood of photon interactions both within the patient and the detector, thus reducing the relative Poisson noise in the images. Furthermore, the proportion of interactions which are photoelectric absorption increases, thereby increasing the differential attenuation of different materials in the field and improving intrinsic subject contrast. Utilising the lowest available beam energy therefore maximises contrast and minimises noise at any given dose level[123]. For aS500 EPIDs operating under the IAS2 'Sync' mode the lower the dose-rate the more lines which can be read out in between beam pulses, the lower the dose per frame and thus the more flexibility in dose which can be specified per image. Even for aS500-II or aS1000 EPIDs operating under 'Rad Shot' mode it is attractive to acquire images at a low dose-rate for safety reasons: should the

beam not terminate as expected once an image has been acquired then a lower dose-rate will result in less additional dose being delivered to the patient whilst the operators recognise an equipment failure and manually intervene to stop the beam. All linacs considered in this study therefore have either two or three clinical beams for imaging: one or two (depending upon the linac technology) which are the standard beams for treatment, plus one which is used purely for imaging. Table 3.2 includes the treatment and imaging dose-rates for all linacs considered in this study.

Another consideration when acquiring images is whether to use standard quality (SQ) or high quality (HQ) mode. In both centres all dedicated imaging beams are routinely acquired using the SQ mode so that the concomitant dose from imaging is minimised. If the quality of an image is poor – for example, when imaging a particularly large patient so that the signal at the detector and thus the CNR are low – the decision may be taken to increase the dose in an effort to boost image quality by utilising HQ mode for subsequent images. However, when images are acquired during treatment field delivery the HQ mode is always used because extending the portion of the field which contributes to imaging by increasing the number of frame averages has no bearing on the patient dose burden.

Experiments with the QC-3V phantom were performed using the aS500 units in Edinburgh, and the PTW phantom was examined using a limited number of aS500 units in Edinburgh and aS1000 units in Oxford. Images of the QEPI1 phantom were acquired using standard clinical imaging modes on all models of EPID and all available linear accelerators. This enabled the performance comparison of the different generations of equipment presented in section 3.8.

Two linacs in table 3.1 are identified as being equipped with Varian's On-Board Imaging (OBI) system. This comprises a kilovoltage X-ray tube and flat panel detector which together provide a diagnostic grade imaging beamline at 90° to the megavoltage treatment beam. OBI equipped linacs are capable of acquiring kilovoltage energy radiographs, fluoroscopy sequences and volumetric cone-beam CT (CBCT) scans of the patient in the treatment position. Performance evaluation of OBI is considered in detail in chapters 4 and 5.

3.3. QC-3V Phantom

3.3.1. Phantom Introduction

A photograph of the QC-3V phantom is shown in figure 3.1 alongside a schematic diagram of its components. It is based on work by Rajapakshe and Shalev[281] and was one of the first phantoms for EPID performance evaluation[123, 135].

Constructed from layers of metal and plastic, the QC-3V phantom consists of three major parts:

- **N1–N4**. The numbers 1 to 4 machined into 15 mm thick lead blocks at progressive increasing depths of 1 to 4 mm. These allow a subjective check of image contrast.
- **B1–B5**. Alternating strips of 15 mm lead and PVC septa, forming bar patterns increasing in spatial frequency through 0.10, 0.20, 0.25, 0.45 and 0.76 cycles / mm in the order indicated in the diagram. An MTF can be extracted using regions of interest placed on the bar patterns, using the methods described in section B.17.2.
- **U1–U6.** Uniform patches of different thicknesses of lead, aluminium and PVC, increasing in attenuation in the order listed.

Complementing the phantom is the PIPSPro image analysis software automatically processes its images and generates summary indices of image quality[372, 381]. Version 3.2.2 of the software was used in this work.

A comprehensive IQWorks analysis tree was developed to process images of the QC-3V phantom and calculate the MTF curve, contrast-to-noise ratio (CNR) and the mean signal under each of the patches U1–U6. In line with the approach taken by PIPS Pro, the MTF was calculated using the basic bar pattern algorithm described in section B.17.2, with no aliasing correction, and the tree was optimised to accept a pair of images acquired in quick succession under the same conditions, which should therefore be identical except for the influence of stochastic noise. The difference image was used in the MTF calculation to compensate for the presence of noise fluctuations.

3.3.2. Comparison of IQWorks against PIPS Pro

Images were acquired of the QC-3V phantom using a Varian 21EX linear accelerator in Edinburgh equipped with a Varian aS500 amorphous silicon EPID

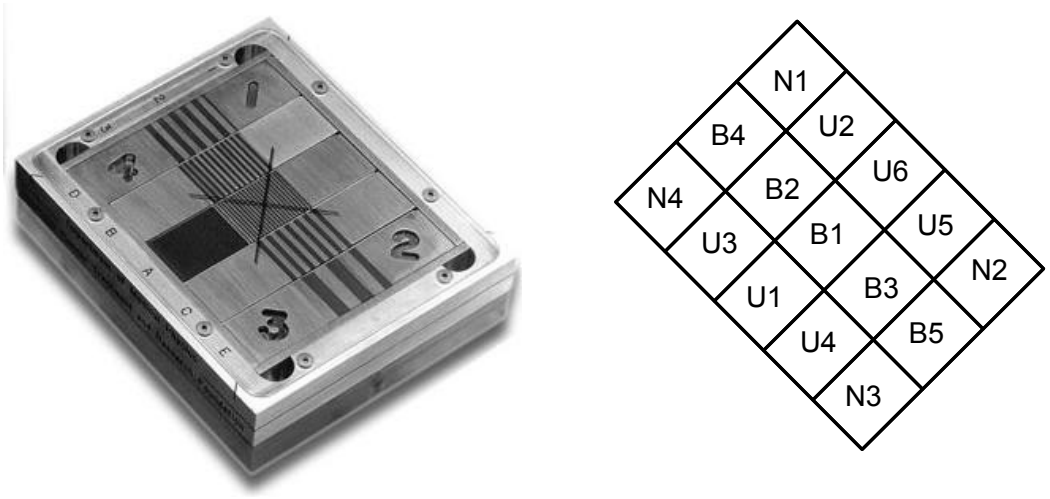


Figure 3.1. – Photograph and schematic diagram of the QC-3V phantom. Descriptions of the phantom's components are included in the text.

(LA4 in table 3.2). In this evaluation a treatment imaging beam of energy 8 MV and dose-rate 500 MU/min was utilised and images were acquired using HQ mode (taking an average of 10 frames and delivering a dose of 6 MU / image). Two images were acquired of the phantom at isocentre, representing performance mid-plane in the patient, and two images with the phantom on the surface of the detector, so that the potential for evaluating intrinsic detector performance could be investigated. Each time, the imager was centred on the beam central-axis and at a SSD of 140 cm.

MTF curves calculated by PIPS Pro and IQWorks for both scenarios are plotted in figures 3.2 and 3.3. Taking PIPS Pro as the 'gold standard', the deviations of IQWorks from the PIPS Pro measurements are also plotted in the figures.

Whilst considering the use of the QC-3V phantom and developing the 'Bar Pattern MTF' IQWorks modules it was suspected that the noise correction may not actually be required. It is inconvenient to always need to acquire two sets of images in any experiment so that a difference image can be calculated. It would therefore be advantageous if the noise correction could be avoided and the number of images thus required reduced by half. To investigate this each image was also analysed individually by IQWorks, using an alternative noise correction strategy (average standard deviation within uniform ROIs), then a multi-layer image was analysed with no noise correction. In addition, for comparative purposes MTF curves with no noise correction were manually constructed from intermediate information written to disk by PIPS Pro during its more detailed calculation. These results are also included in figures 3.2 and

3.3. It is noted that, apart from the points for the highest spatial frequency bar pattern, all curves in the figures essentially overlap, making it very difficult to distinguish between them and indicating excellent agreement between PIPS Pro and IQWorks.

For the phantom at isocentre, the noise-corrected IQWorks curve agrees with that of PIPS Pro to within 1% across all spatial frequencies, and the curve for the phantom on the surface of the detector agrees within 2%. It is thought that this increased discrepancy is a geometrical magnification effect: when placed on the detector the phantom image covers less of the detector's surface than when it is projected from the isocentre plane. The number of pixels across each object within the phantom image is therefore smaller, making it more difficult to place ROIs in exactly the same locations in both software packages and increasing the level of experimental uncertainty. From these results IQWorks can be considered able to successfully mimic the behaviour of PIPS Pro in calculating MTF.

All investigated strategies in the noise correction experiment – including applying no noise correction – also agree with the standard PIPS Pro analysis to within 1% when the phantom is at isocentre and within 2% when on the surface of the detector, except for at the highest spatial frequency, where the difference is considerably worse. Because the highest frequency bar pattern has the lowest modulation it is logical that this is the one most influenced by the presence of noise. It can therefore be concluded that to generate an MTF curve consistent with that produced by PIPS Pro it is important to always acquire a pair of images and include a difference image noise correction. However, given the inherent uncertainty in the algorithm due to the lack of the standard Coltman anti-aliasing correction it is arguable that, if there were no requirement to be directly compatible with PIPS Pro, the noise correction is an unnecessary refinement. It is suggested that for regular QC checks, and during optimisation exercises, it is justifiable to analyse only individual images rather than pairs, with any repeated images being used to assess the overall uncertainty in the experimental technique rather than directly for noise correction.

In PIPS Pro the MTF is normalised to unity at the modulation of the lowest spatial frequency bar pattern: 0.10 lp/mm. Performing a similar normalisation in IQWorks, all of the methods yield the same f_{50} regardless of the noise correction strategy applied: 0.44 lp/mm for the phantom at isocentre, and 0.37 lp/mm for the phantom at the surface of the detector. A degradation in MTF as the phantom moves towards the detector is expected due to the magnified

bar patterns covering more detector pixels, thus having lower effective spatial frequencies and being easier to resolve. The reproducibility of the f_{50} metric across all noise correction strategies reinforces the finding that noise correction is not essential.

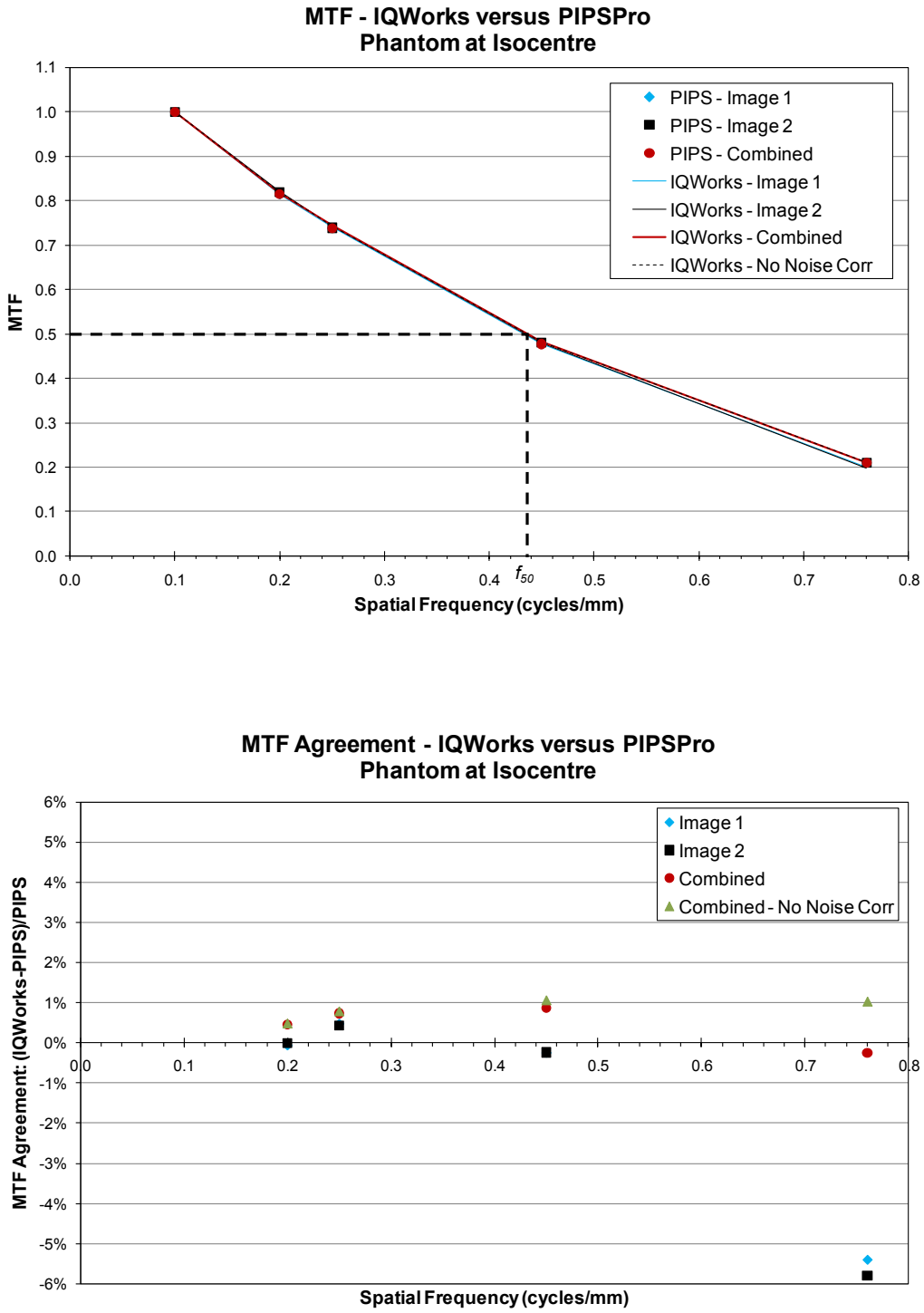


Figure 3.2. – Comparison of the modulation transfer function (MTF) results obtained when analysing images of the QC-3V phantom positioned at isocentre. Top–Calculated MTF curves. Bottom–Discrepancy between IQWorks and PIPSPro.

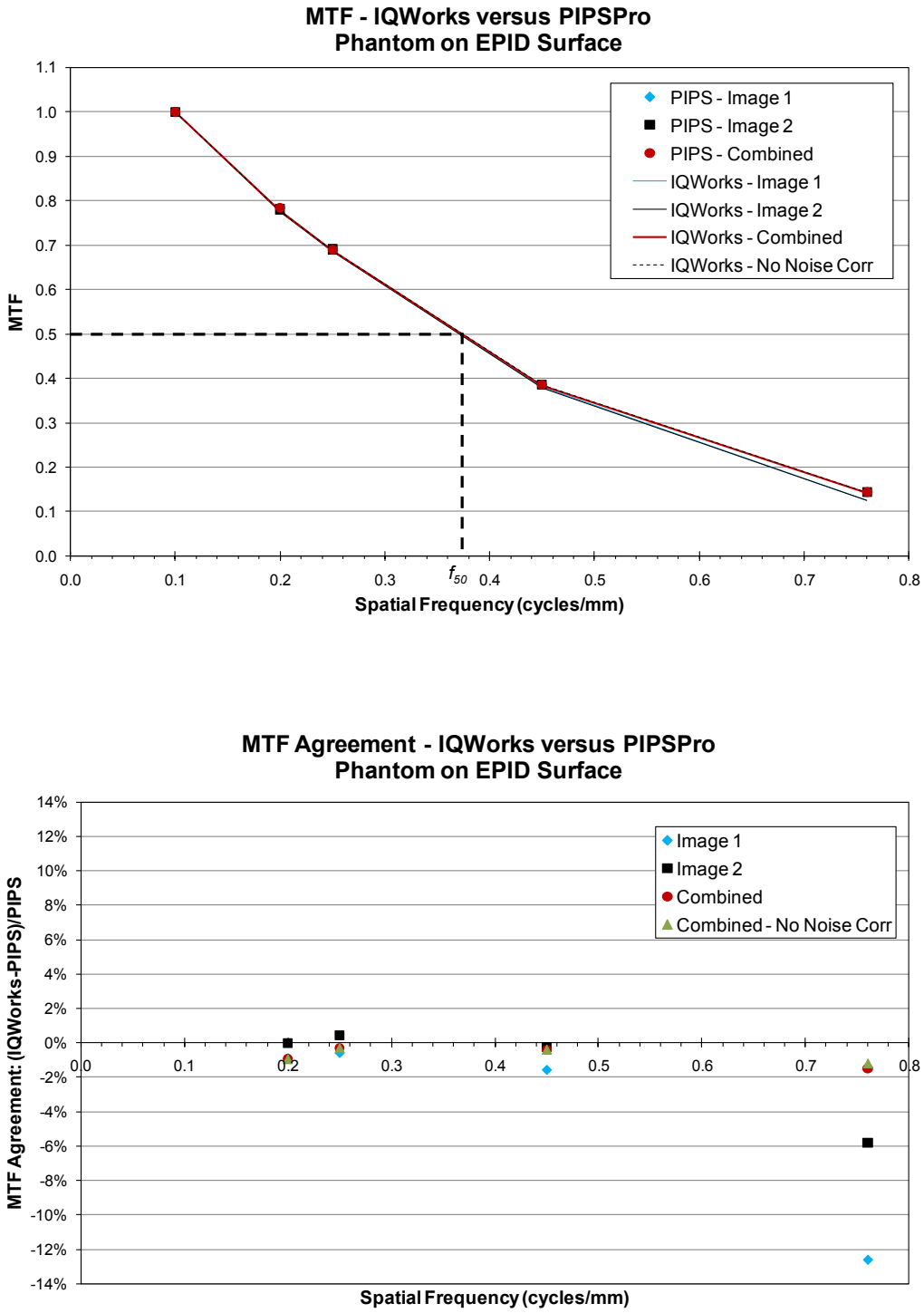


Figure 3.3. – Comparison of the modulation transfer function (MTF) results obtained when analysing images of the QC-3V phantom positioned on the surface of an EPID at 140 cm SID. Top–Calculated MTF curves. Bottom–Discrepancy between IQWorks and PIPSPro.

For the purposes of validating IQWork statistical algorithms, the mean and standard deviation in the pixel values in each uniform ROI were calculated and compared with those produced by PIPS Pro. From the results in figure 3.4 it is evident that IQWorks and PIPS Pro agree in all ROIs on both images to within one standard deviation, the experimental uncertainty. Agreement between packages for the standard deviation (SD) measurements was also good - to within 10% for ROIs 1–4 and within 20% for ROI 6. Greater variability is expected between measured SDs than for mean values because these are more sensitive to the exact pixel values sampled within the ROIs, and the absolute numerical values are significantly smaller (of the order of 10 for SD rather than 10^3 for means) so that a small numerical difference in SD measurement appears proportionately larger. Furthermore, the result that the SD discrepancy is worse for ROI 6 is also as expected because this is the most attenuating patch of lead in the phantom, under which the number of photons reaching the detector is least and therefore the relative noise is greatest. Only the results for the phantom at isocentre are included here because this is a test of basic statistical calculations. However, those for the phantom on the surface of the detector also agreed well between the two packages.

It is noted that the standard deviation measured by IQWorks in ROI 5 is significantly lower than that of PIPSPro, with agreement between the packages being considerably poorer for this ROI than for all the others. This is because the QC-3V phantom used in the experiment had a flaw in the material forming the block U5, introducing a non-uniformity. The IQWorks analysis tree was modified to compensate for this, whereas the PIPS Pro algorithm assumes that the whole area of U5 is available for analysis. This is illustrated in figure 3.5 and demonstrates the flexibility of IQWorks over commercial packages in dealing with unexpected test conditions.

CNR was calculated according to equation 2.5. In the literature describing the QC-3V phantom[281] the noise level σ is defined as

$$\sigma = \sigma_{sub}/\sqrt{2} \tag{3.1}$$

where σ_{sub} is the standard deviation over the highest frequency bar pattern in the difference image. However, it was discovered that both PIPS Pro and the Excel macro bundled with the software take σ_{sub} as the mean standard deviation across all bar patterns and the $\sqrt{2}$ factor is ommitted. When the same approach is followed in IQWorks, the CNR results agree with those of PIPS Pro to better than 2%: 530 ± 10 for the phantom at isocentre, and 495 ± 10 for the phantom on

the surface of the detector. Nevertheless, the difference between the description in the literature and the algorithmic implementation in the commercial software is an important result which is not clearly documented.

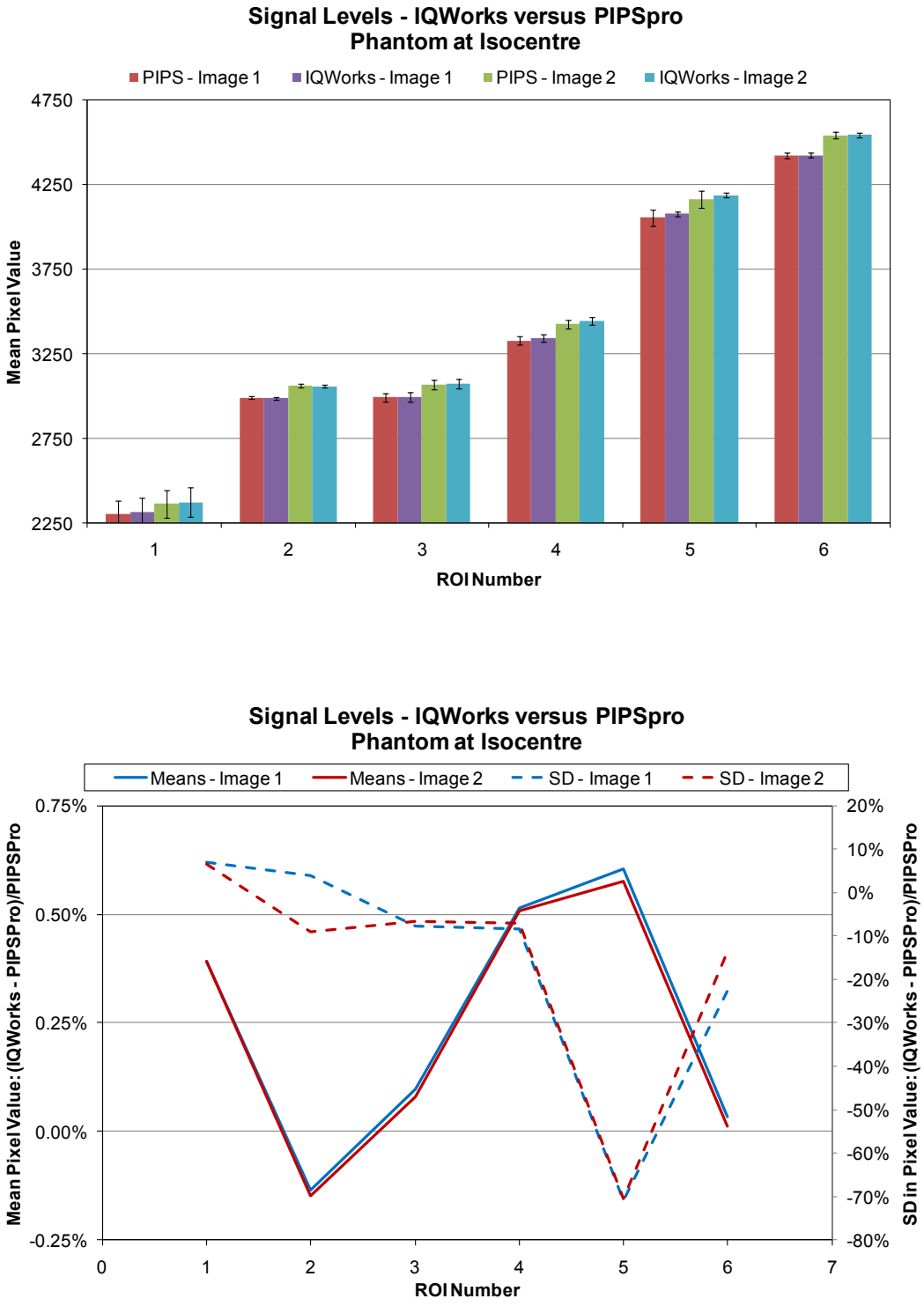


Figure 3.4. – Comparison of the mean and standard deviation in pixel values calculated by IQWorks and PIPSPro in regions of interest (ROIs) placed on the uniform patches U1 to U6. Top–Side by side comparison of results, where the error bars indicate ±1 standard deviation (SD). Bottom–Difference between IQWorks and PIPSPro, as a fraction of the PIPSPro result. Solid lines represent agreement between the mean values within ROIs and the dashed lines are for the SDs.

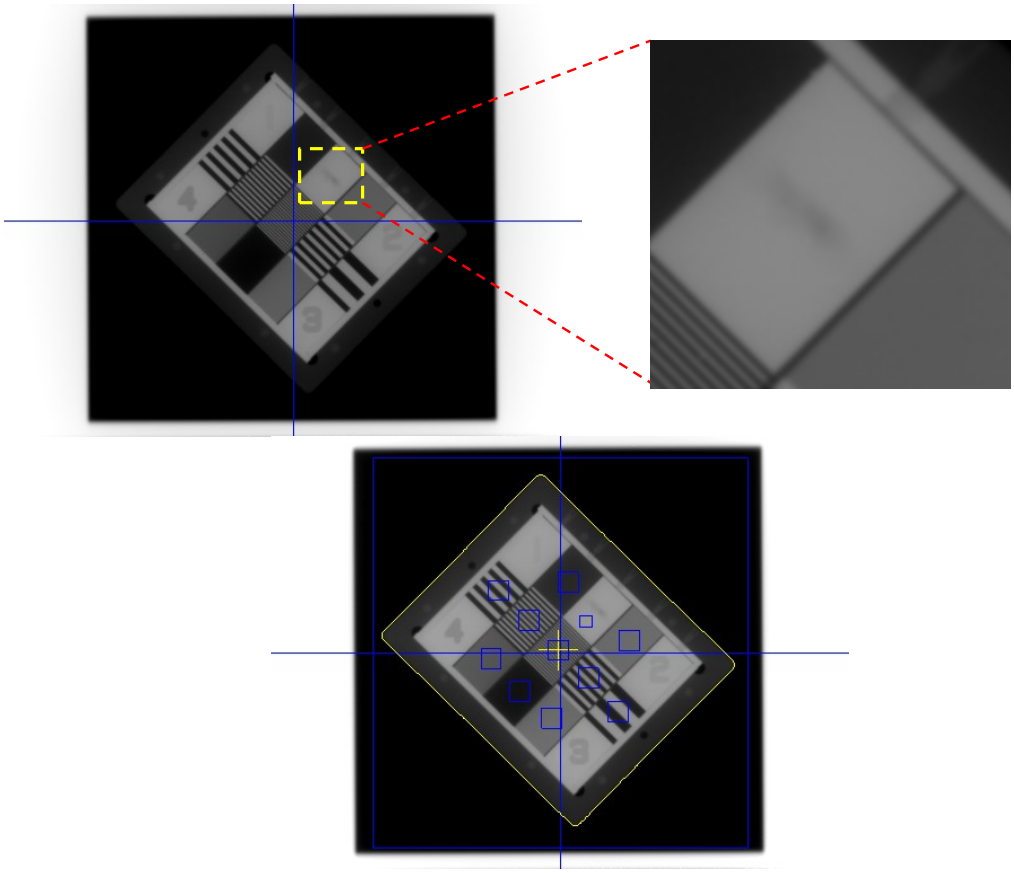


Figure 3.5. – Customising the IQWorks analysis tree to compensate for a flaw in a particular QC-3V phantom. Top: 6 MV EPI of the phantom with the flawed U5 region indicated and blown up. Bottom: Placement of ROI in IQWorks analysis tree modified to avoid flawed region.

3.4. PTW EPID QC Phantom

3.4.1. Phantom Introduction

Intended to be a comprehensive phantom for EPID performance evaluation, the PTW 'EPID QC' phantom contains similar analysis components to the QC-3V as well as some additional features[275, 370]. A photograph of the internal configuration of the phantom is shown in figure 3.6 and a schematic diagram of these in figure 3.7.

Analysis components include:

- **S1–S10.** Copper step wedge, with steps of thickness 26.7, 19.6, 16.6, 13.5, 11.1, 8.6, 6.3, 4.1, 2.0 and 0.0 mm. These nominally result in transmission of a 6 MV photon beam of 50% to 100%, in 5% increments.
- **G1–G16.** Brass blocks of cross-section 10 mm $\times$ 10 mm and height 19.6 mm. G1–G8 are arranged vertically and G9–G16 horizontally, with a nominal centre-to-centre block spacing of 20 mm.
- **B1–B18.** Bar patterns comprising brass bars of approximately 20 mm in height separated by air gaps of equal width. Patterns B1–B4 are angled with respect to the axes of the image matrix and have spatial frequencies of 0.125, 0.167, 0.25 and 0.33 cycles / mm. B5–B11 are aligned with the Y axis and B12–B18 with the horizontal axis, and the patterns in each of these groups have spatial frequencies of 0.5, 0.67, 1.0, 0.59, 2.0, 2.5 and 3.33 cycles / mm.
- **Ringed Blocks ABCD.** Four brass bars cross-section 10 mm $\times$ 10 mm and heights A: 19.6 mm, B: 8.6 mm, C: 13.3 mm and D: 4.1 mm. These are intended to result in 10%, 30%, 20% and 40% transmission of a 6 MV photon beam. Located at the four corners of the phantom and in two positions along a diagonal, they are used to assess local signal linearity. The circular brass rings surrounding the ABCD blocks may be used to aid phantom alignment.
- **CD1.** Contrast-detail test object consisting of a 20 mm aluminium block in which is drilled holes aligned in rows of increasing depth and columns of decreasing diameter, when moving towards the centre of the phantom.

A key difference between the EPID QC and QC-3V phantoms is that the test components within the QC-3V are angled so that their edges align with diverging rays at a distance of 100 cm from the X-ray source. Whilst this minimises

penumbra due to rays passing through test components at oblique angles, thus allowing the test objects to be sufficiently thick to have high attenuations at megavoltage energies, there are a number of issues with this approach. Firstly, it has the side-effect that the phantom can only be positioned at isocentre, thereby preventing investigation of intrinsic detector performance by a phantom placed on the surface of the detector, unless the detector itself is raised to 100 cm SSD, something not possible for all EPIDs. In addition, the focusing of the test components is achieved by mounting them on an angled plastic base. Therefore, rather than the components themselves being intrinsically focused each is only grossly angled back towards the focal spot. For example, the bars of patterns B1–B18 are not focused although each block of line-pairs is angled as a whole. This somewhat defeats the original objective of focusing. Furthermore, because focusing assumes the centre of the phantom is aligned with the central axis it cannot be repositioned to examine image quality in different locations across the field of view.

Another limitation of the PTW EPID QC phantom is that it relatively large, being a square of side 25 cm. Unfortunately, because the epidSoft package requires the whole phantom to be in the FOV so that it can localise landmarks and place analysis ROIs, this limits the acceptable range of EPID vertical positions. For Varian detectors it was found that the EPID could at maximum be 20 cm beyond isocentre before the critical alignment landmarks began to fall outside the FOV. This is a particular limitation because patient imaging geometries tend to require the EPID to be at least 30 cm beyond isocentre, meaning that the EPID QC phantom can never be utilised under truly representative clinical imaging conditions.

In contrast, the less complex QC-3V phantom places almost no restriction on the location of the detector and can be positioned anywhere in the field of view.

Apart from the CD1 component, which to an extent can be interpreted visually, the designers of the PTW EPID QC phantom intend that images of the phantom are processed by the dedicated PTW 'epidSoft' analysis package. Experiments were performed to determine whether IQWorks could be configured both to replicate and extend the capabilities of epidSoft.

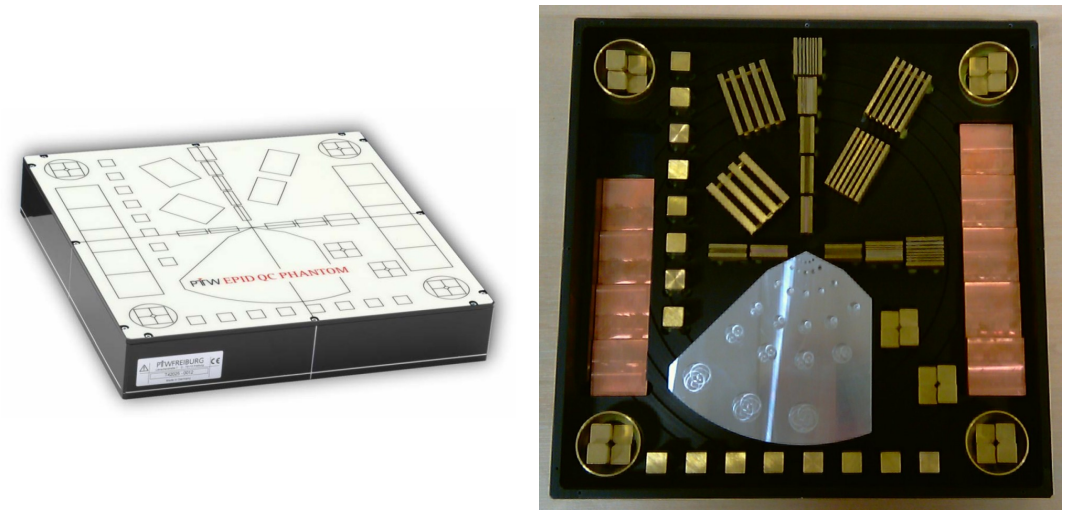


Figure 3.6. – Photographs of the PTW EPID QC phantom. Left: external surface of the phantom showing laser alignment lines and the location of internal test objects. Right: Photograph of the interior of the phantom showing brass and aluminium test objects mounted on an angled plastic base.

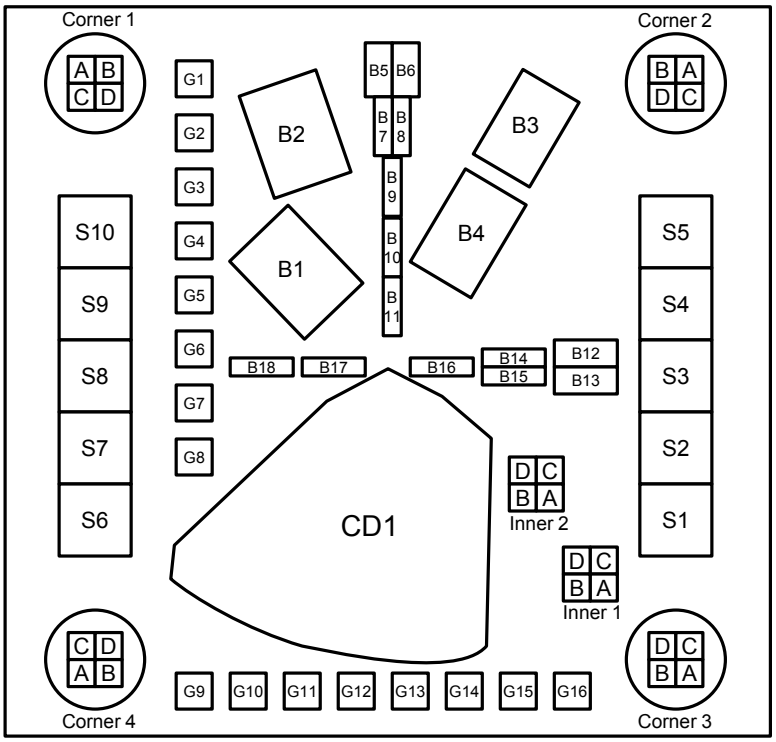


Figure 3.7. – Schematic diagram of the PTW EPID QC Phantom with all internal test objects labelled. These are described in more detail in the main text.

3.4.2. MTF: Comparison of IQWorks with PTW epidSoft

epidSoft's MTF calculation is performed on single images in isolation and involves no corrections for noise. Although it is based on the standard Coltman formalism described in section B.17.2 it was discovered that the aliasing correction described in the epidSoft manual[275] is subtly different from that in equation 3.3. Whereas the sign of successive terms in equation 3.3 alternates from positive to negative, the correction in the epidSoft manual involves only additive corrections. i.e. instead of the correction being

$$MTF'(f) = MTF(f) + \frac{MTF(3f)}{3} - \frac{MTF(5f)}{5} + \frac{MTF(7f)}{7} - \dots \quad (3.2)$$

it is

$$MTF'(f) = MTF(f) + \frac{MTF(3f)}{3} + \frac{MTF(5f)}{5} + \frac{MTF(7f)}{7} + \dots \quad (3.3)$$

It is unclear whether this difference is intentional or in error. However, an additional noise correction matching that of equation 3.3 was incorporated into the 'Bar Pattern MTF' analysis module to accommodate it.

Two images of the phantom were analysed by both epidSoft and IQWorks - one acquired using SQ mode at 8 MV, 100 MU/min (4 frame averages, at a dose of 3 MU) and the other using HQ mode at 15 MV, 500 MU/min (10 frame averages, at a dose of 8 MU). Due to the differences in energy it was expected that there should be a significant difference in the image quality metrics calculated from the two images.

Surprisingly, the agreement between the MTF curves calculated by the two packages was considerably worse than between IQWorks and PIPS Pro for the QC-3V phantom. A first observation was that epidSoft forces the modulation from all bar patterns to zero after the Nyquist frequency, significantly altering the nature of the curve at higher spatial frequencies. This behaviour was incorporated into the epidSoft aliasing correction implemented by IQWorks. However, even following this adjustment, the agreement between IQWorks and epidSoft was of the order of 10% at 8 MV and 15% at 15 MV, as shown in figures 3.8 and 3.9. Interestingly, for both energies the error gradually increases negatively with spatial frequency. It was suspected the aliasing correction might be at fault so the IQWorks analysis was repeated both with no corrections and with the standard Coltman corrections. Using no correction clearly overestimated the MTF at all frequencies, but the Coltman correction actually improved

agreement to within 10% at 15 MV and with no evidence of a systematic trend in error with spatial frequency.

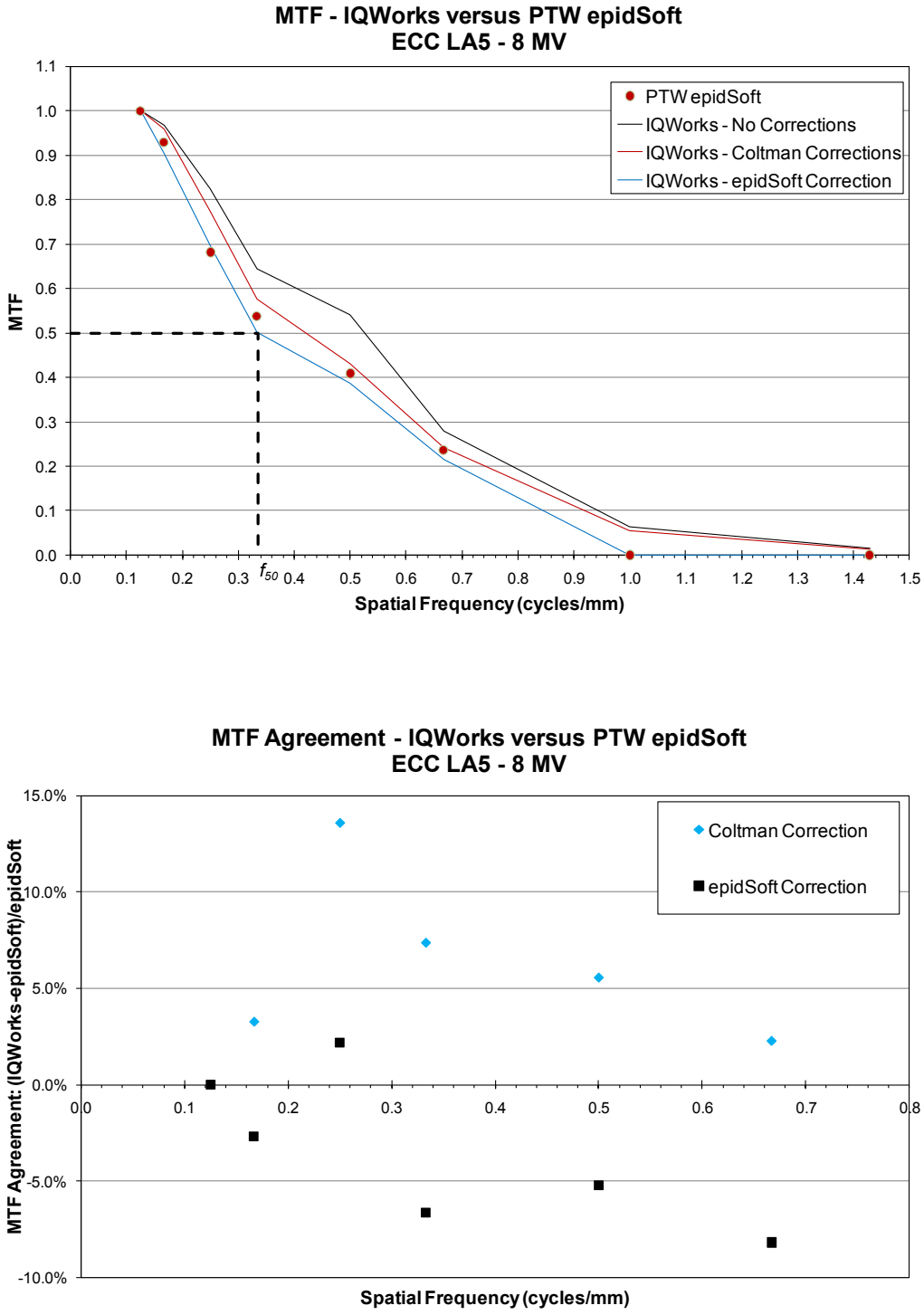


Figure 3.8. – Comparison of the modulation transfer function (MTF) curves calculated by IQWorks and epidSoft from the angled (B1–B4) and vertical (B5–B11) bar patterns in the PTW EPID QC Phantom. The image was acquired at 8 MV, 100 MU/min. Top–MTF Curves. Bottom–Discrepancy between IQWorks and epidSoft for the two anti-aliasing options implemented in IQWorks.

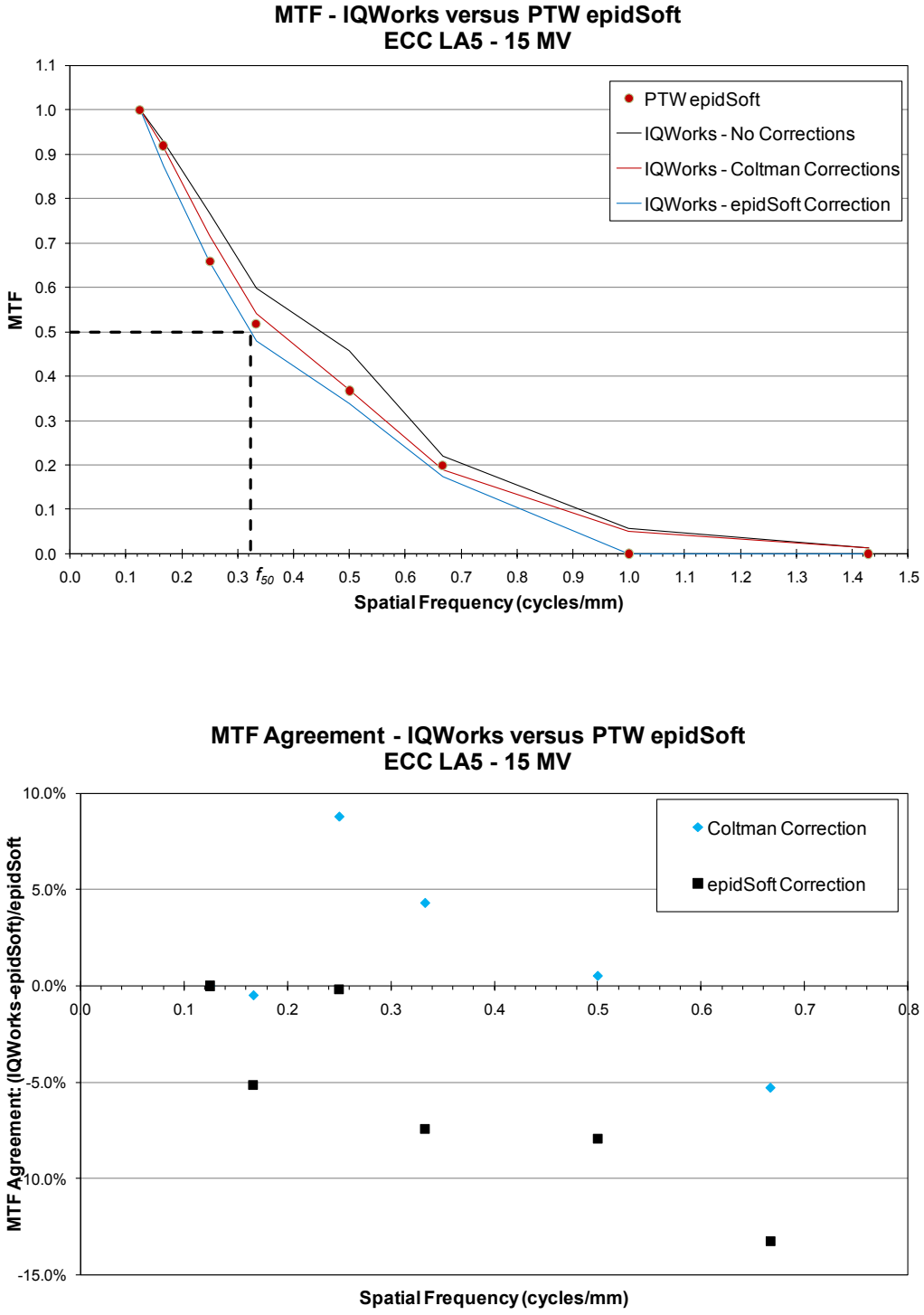


Figure 3.9. – Comparison of the modulation transfer function (MTF) curves calculated by IQWorks and epidSoft from the angled (B1–B4) and vertical (B5–B11) bar patterns in the PTW EPID QC Phantom. The image was acquired at 15 MV, 500 MU/min. Top–MTF Curves. Bottom–Discrepancy between IQWorks and epidSoft for the two anti-aliasing options implemented in IQWorks.

Examining the regions of interest employed in the analysis by epidSoft it was apparent that these seemed rather large in relation to the area of the bar patterns in the image (see screenshot in figure 3.10) and it was suspected that phantom alignment, and the subsequent accuracy of ROI placement, might have an influence on the MTF. Using the same analysis routines as previously (and applying the epidSoft correction in IQWorks), MTF curves were calculated for a new image, this time acquired using a higher resolution aS1000 detector under Rad Shot mode at 6 MV, 100 MU/min and at a dose of 1 MU. (This new experiment was performed in Oxford whereas the previous evaluation had been undertaken in Edinburgh. Detailed comparisons between the packages were only possible when the author had moved to Oxford and access to the Edinburgh linacs was no longer available. However, it was initially thought that the disagreement between the packages might be the result of operator error during the earlier experiments, so repeating the measurements in Oxford was seen as an independent check. Furthermore, it was hoped that any beam stability artefacts which might be influencing results would be removed by the newer the Rad Shot mode available on the systems in Oxford[343, 347].) The analysis was repeated multiple times, but each time with the ROIs being shifted systematically over a 9×9 grid $\pm$ 1 pixel either side of the original location in all directions. This resulted in 9 MTF curves in total (the original, plus 8 curves from slightly mis-aligned ROIs) and the mean curve calculated by each package, along with an envelope indicating the minimum and maximum modulations at each spatial frequency, are plotted in figure 3.11.

From the results it is clear that both packages are sensitive to accurate identification of the phantom origin in the software before processing images, especially at the higher spatial frequencies where the bar patterns are smaller and hence easier to miss if the ROI is placed incorrectly. epidSoft demonstrates considerably greater variability, grossly overestimating the MTF by 150% at the highest spatial frequency, and it is suspected that this is because it employs larger ROIs than IQWorks that are intended to extend fully across the test patterns, as implied in the screenshot in figure 3.10. When these ROIs are misaligned, even by a small amount, they extend beyond the edges of the test patterns and the variance within them increases substantially. Errors are biased towards overestimation of the MTF, which is logical because the standard deviation in an ROI will be greater if the ROI extends to other parts of the phantom. Although ROI placement issues appear to contribute to the lack of agreement between the two packages, the effect on the relatively large-area low-frequency bar patterns

is far less pronounced than at higher frequencies and is insufficient for this to be the only source of error.

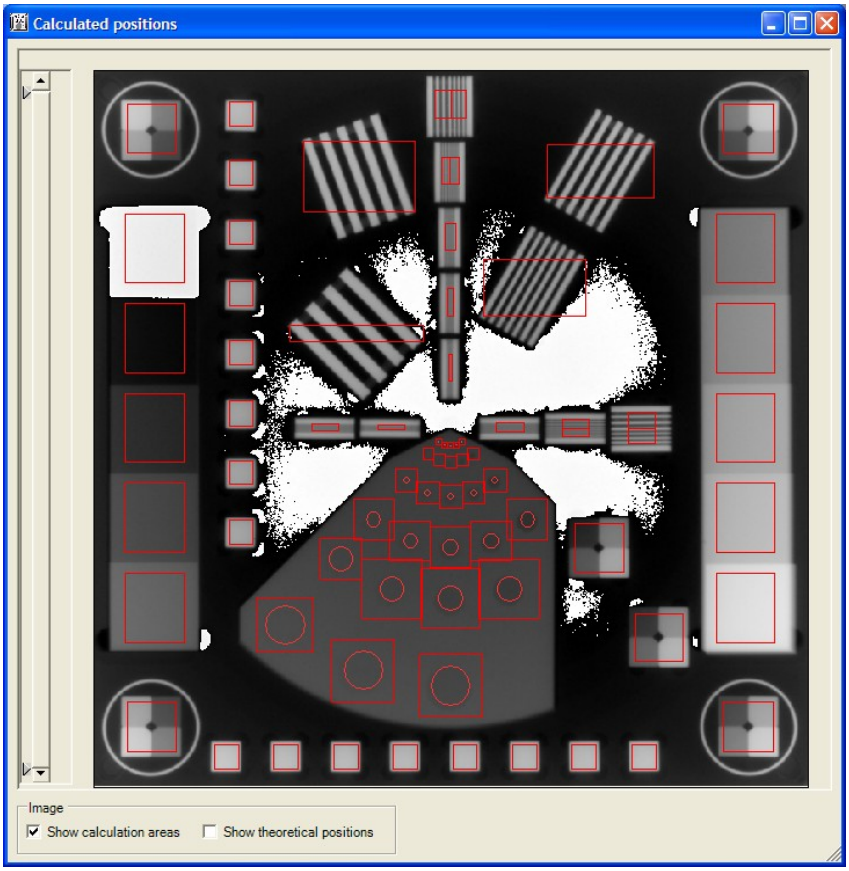


Figure 3.10. – Regions of interest used by PTW epidSoft.

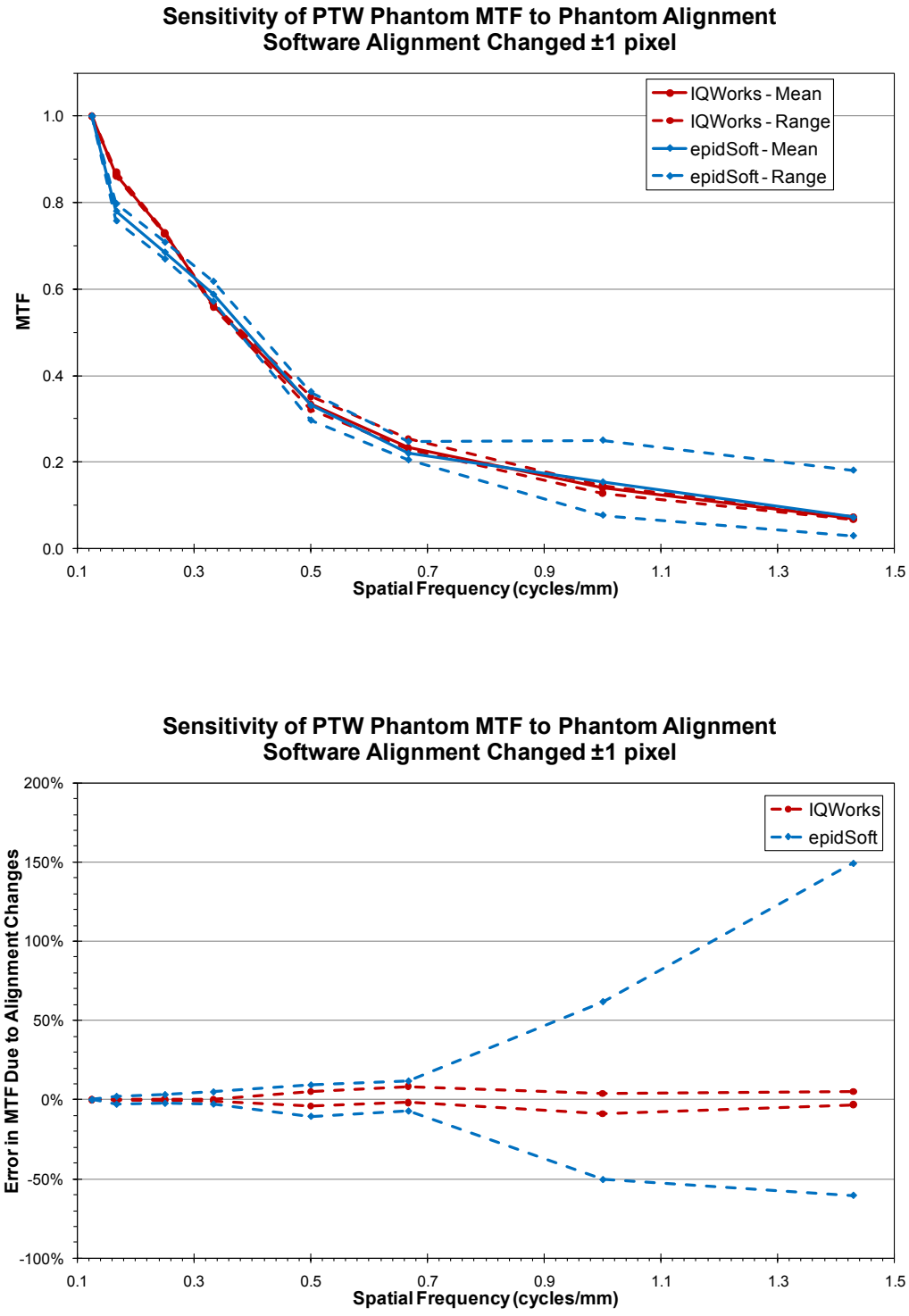


Figure 3.11. – Sensitivity of PTW EPID Phantom MTF to alignment of phantom in software. Comparison of results from IQWorks and PTW epidSoft when phantom origin is shifted by ±1 pixel in each direction in the software. Top–Modulation transfer function (MTF) curves. Bottom–Range of possible errors.

Another experiment was performed, but this time with the MTF curves being compared between 10 separate images acquired under the same acquisition settings as previously. This time the ROIs were placed as accurately as possible on each image but the phantom was physically removed from the couch then repositioned and realigned, again as accurately as possible, between each acquisition. The results are plotted in figure 3.12. Errors introduced through repeatedly positioning of the phantom are similar between both packages, with the magnitude being considerable (>10%) for spatial frequencies > 0.7 cycles/mm. It is thought this is due to two factors: the physical size of the gross bar pattern test objects being smaller as the spatial frequency increases, making them more sensitive to a geometric miss, and the higher frequency patterns being aligned with the pixel matrix being susceptible to interference artefacts which change significantly with slight translations or rotations of the phantom. However, as before the low-frequency bar patterns are relatively tolerant of alignment changes, so further investigation was required to identify the significant discrepancy between IQWorks and epidSoft at the lowest spatial frequencies.

Other workers have described issues with epidSoft's file-handling, reporting problems with 16-bit DICOM files and asserting that epidSoft cannot interpret these properly[204, 269]. In the examples in the literature this has caused discrepancies in CNR and other signal-based image quality metrics but there has been no description of how this influences MTF calculations. The same image used to generate the results in figure 3.11 was loaded by IQWorks and saved back to disk in a range of different file formats.

These included:

1. 8-bit bitmaps, in which the pixel values were mapped into the new range 0–255 by rescaling.
2. Matrix of 16-bit unsigned integers.
3. Matrix of 16-bit signed integers.
4. 1–3 loaded back into IQWorks, inverted, then saved as the same file format.
5. The original DICOM image loaded into IQWorks, inverted, then saved as 1–3.

Each of the new images was loaded into IQWorks and the MTF calculated using automatic placement of ROIs. Despite the manipulation of the pixel values all

MTF curves agreed to 3 decimal places across all spatial-frequencies. This was as expected: the MTF curves are formed from measurements of relative variance which should be insensitive to the scaling and sign of the pixel values within each ROI, as long as the data covered by each ROI have been manipulated in the same way.

However, when the images were loaded into epidSoft differences in image-handling were evident from the considerably different visual representations of the images. This is illustrated for a selection of the images by the screenshots in figure 3.13, in which epidSoft and IQWorks representations are displayed alongside each other for the purposes of comparison. The default file format for Varian EPID images is 16-bit DICOM and the images on the left in the figure indicate how this is interpreted by two packages. It is immediately apparent that the epidSoft image appears inverted with respect to that in IQWorks. However, with reference to the schematic diagram of the phantom in figure 3.7, it is surprising that the top left and bottom right steps of the copper wedges (steps S10 and S1 respectively) both appear black in the epidSoft image, because these correspond to the lowest and highest attenuation steps in the phantom. Attenuation increases monotonically between each pair of adjacent steps in the wedge so that one would expect a gradual change in grey level when moving from one step to the next, and this is what is observed in the IQWorks image. However, in the epidSoft image there appears to be a shift from very high signal to very low signal between steps S9 and S10. Furthermore, in other regions of high signal, such as next to the bar pattern B4 or the 'Inner2' square block, there are specks of very low signal in the epidSoft image which are not present on the IQWorks image and are definitely not localised regions of high attenuation. These artefacts are what one might expect from wraparound of the pixel values through epidSoft attempting to store them using a numerical representation with insufficient bit depth. As described in section 3.2, the raw pixel values in Varian EPID images arise from 14-bit analogue to digital conversion of the amplified charge in each detector element. They are represented in the DICOM file as 14-bit unsigned integers over the range 0 to 16383 (i.e. not using all of the available 16 bits). Placing ROIs on steps S9 and S10 in IQWorks, the mean pixel values (±1 SD) on each step were measured as 2393 ± 13 and 2615 ± 12 respectively. Because both these values lie between 2^{11} (2048) and 2^{12} (4096) it is unlikely the erroneous representation in epidSoft is due simply to the binary representation of the pixel values: if a binary number was being truncated through an overflow because of insufficient bit depth, or if the most significant

bit was being interpreted as a sign bit, then this would affect pixel values beyond a particular power of two threshold, which is not what is observed here.

Also surprising was the result when the 16-bit DICOM image was mapped to an 8-bit bitmap, as illustrated in the centre pair of images in figure 3.14. These images were formed by rescaling the range of signal values in the original image (0–16383) to the 8-bit range 0–255, with the sign of the pixel values being preserved (i.e. a larger signal being indicated by a more positive number). One would therefore expect the visual representation of the original 16-bit and new 8-bit images, when windowed to show the full range of values, to be very similar. This is what is observed in IQWorks, but the 8-bit image appears inverted in epidSoft. Furthermore, when the 8-bit bitmap image is inverted so that high signal is mapped to low pixel value (i.e. $255 \rightarrow 0$, $254 \rightarrow 1$, etc.), the image appears as expected in IQWorks, and this time the epidSoft and IQWorks representations agree. Inverting the bitmap therefore appears to result in an inverted image in epidSoft, which is encouraging despite the erroneous behaviour when interpretting 16-bit images.

MTF curves from all the images are plotted in figure 3.14, in which it is apparent that no two curves generated by epidSoft are identical. Instead, at least four distinct epidSoft curves lie either side of that produced by IQWorks, with the range being sufficiently large to account for any discrepancy between the two packages. Given the discussion above regarding variance being insensitive to pixel sign (and therefore whether a distribution of values lies at the low or high end of the 8-bit 0–255 range) it is surprising that even the simple 8-bit original and inverted images do not yield similar results in epidSoft.

No satisfactory explanation for epidSoft's behaviour could be arrived at but it is evident there are fundamental issues with how it interprets underlying image data, leading to erroneous MTF calculations. It is therefore doubtful that this package could be utilised to compare results from different vendors' equipment. Furthermore, these flaws will inhibit optimisation exercises (where the range of pixel values will vary between beam energies and EPID acquisition modes) and could potentially result in results appearing different between different revisions of the Varian EPID control software, where – from experience – it is likely there may be subtle changes in storage file format. In contrast, analyses performed by IQWorks are demonstrably more stable and reproducible over a wide range of image encodings.

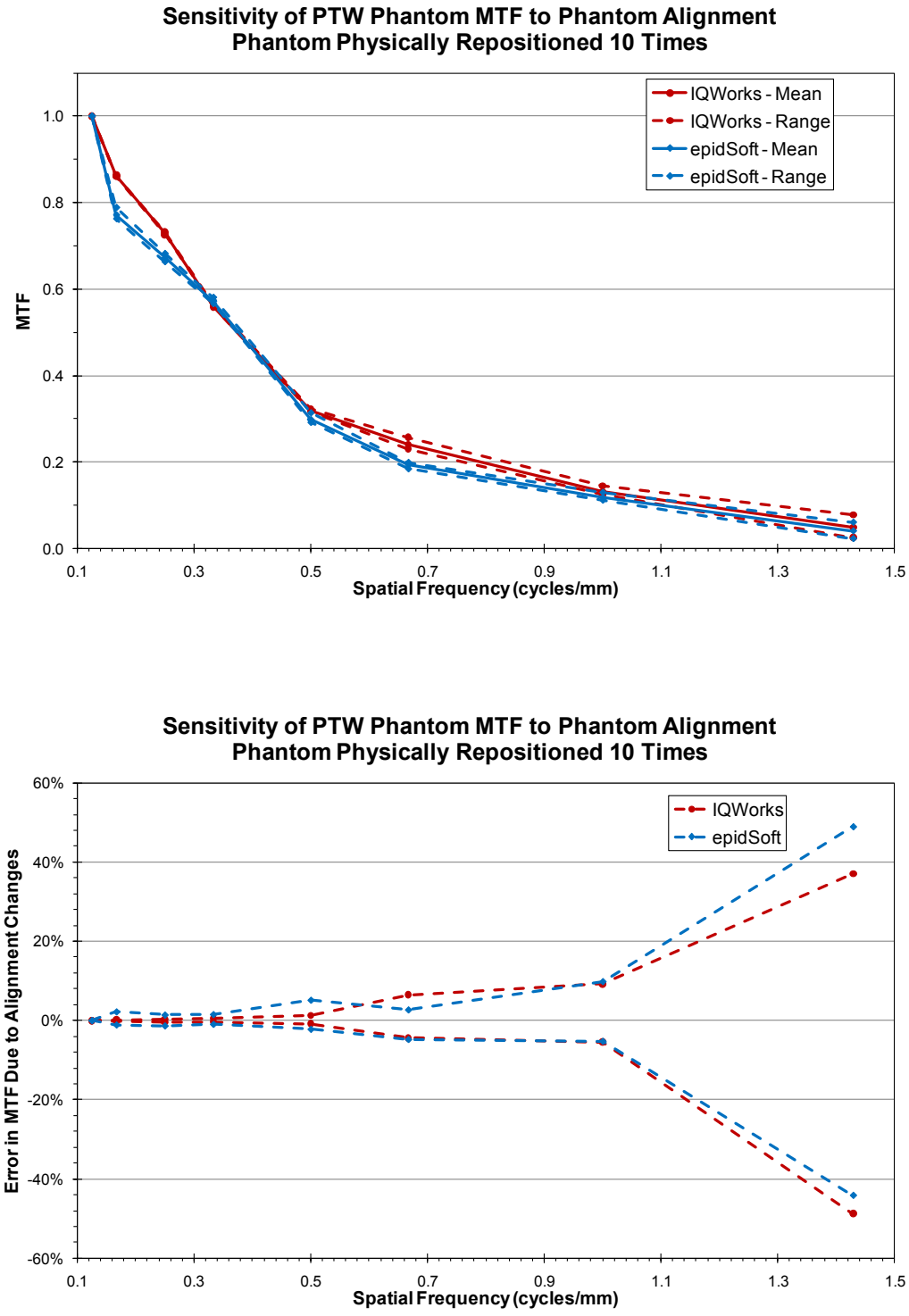


Figure 3.12. – Sensitivity of PTW EPID Phantom MTF to physical alignment of the phantom. Comparison of results from IQWorks and PTW epidSoft when phantom is repeatedly removed then realigned as accurately as possible 10 times. Top–Modulation transfer function (MTF) curves. Bottom–Range of possible errors.

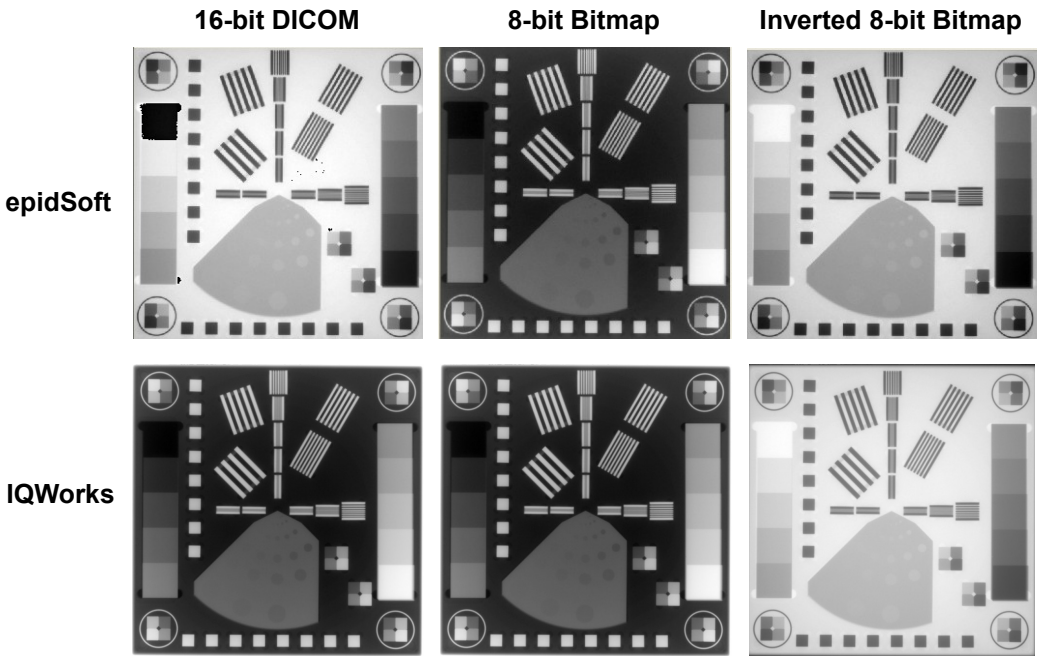


Figure 3.13. – Visual presentation of differently-encoded images by epidSoft and IQWorks. The '16-bit DICOM' image is the default Varian file format for electronic portal images, the '8-bit Bitmap' is this image mapped onto the range 0–255 and the 'Inverted 8-bit Bitmap' is the inverse of this (i.e. pixel values are subtracted from 255). A detailed description of these images is included in the main text.

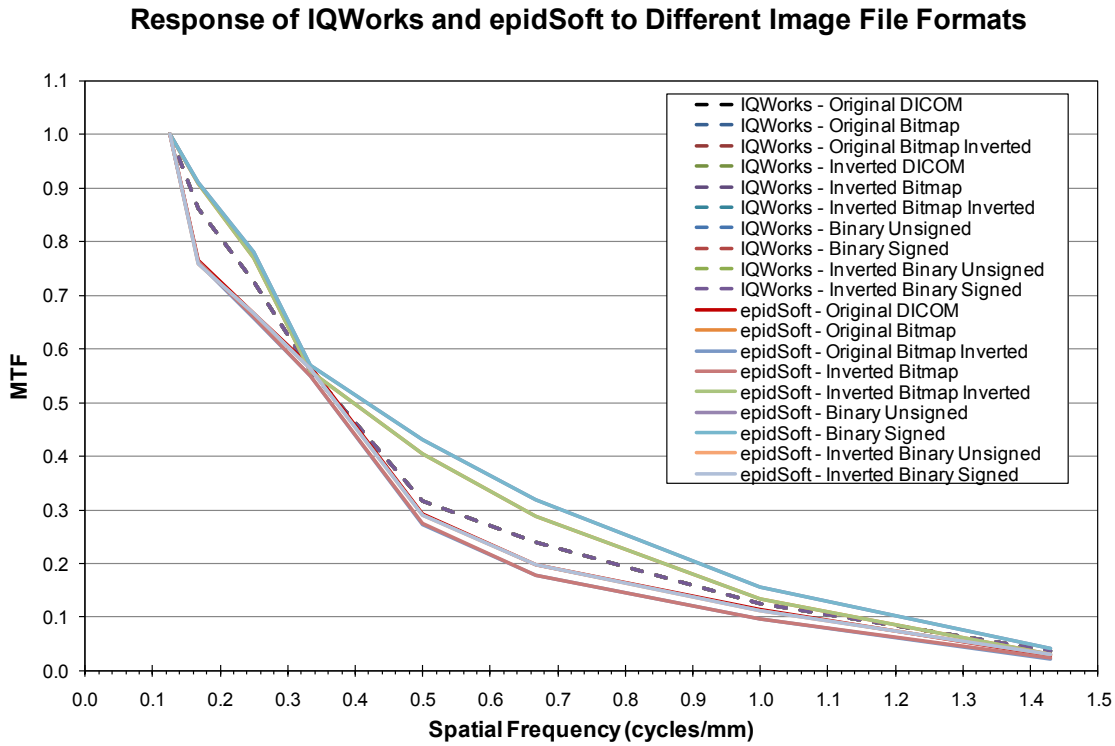


Figure 3.14. – Effect on modulation transfer function (MTF) of submitting the same image to IQWorks and PTW epidSoft using different file formats.

3.4.3. Signal Linearity: Comparison of IQWorks with PTW epidSoft

Signal linearity is calculated by considering the mean pixel values in ROIs located on steps S1–S10 of the copper wedge, with the results being presented as a function of nominal absorption at 6 MV. However, rather than plot absolute signal value epidSoft renormalises and maps the data to a new range so that the reported signal under S10 (100% transmission) is unity and that under S1(50% transmission) is 0.5. i.e. If the measured signal under step i is S_i then the relative signal reported by epidSoft for that step is given by

$$S'_i = 0.5\left(1 + \frac{S_i - S_1}{S_{10} - S_1}\right) \tag{3.4}$$

Images of the EPID QC phantom were acquired at 6, 8 and 15 MV and the linearity curves generated by IQWorks and epidSoft are compared in figure 3.15. The 8 and 15 MV images were acquired using linac LA5 in Edinburgh at a dose-rate of 500 MU/min and dose of 8 MU / image (10 frame averages, HQ Sync mode), whilst the 6 MV images were acquired using linac V3 in Oxford at a dose-rate of 600 MU/min and a dose of 2 MU / image (4 frame averages, HQ Rad Shot mode). In all cases it is evident that the IQWorks results are identical to those of epidSoft, within the limits of experimental error (defined as ±1 SD in the pixel values in the IQWorks ROI at each spatial frequency).

However, a limitation of this analysis methodology is that by remapping the signals to the range 0.5–1.0 the first and last signal points are effectively lost. Furthermore, the validity of the mapping is questionable because the step wedge is only specified to give linear transmission with step size at 6 MV, and then only nominally. One would therefore expect to be able to observe and potentially be able to track over time the relative difference in signal due to energy, yet the curves for all three energies are very similar in the figure.

An alternative approach is to still normalise absolute signal level to unity at nominally 100% transmission but perform no additional range mapping. This generates the curves in figure 3.16, where the differences in energy are far more pronounced. Interestingly, the nominal 50% transmission step actually yields a relative signal of nearer 38% at 6 MV, and the two 15 MV curves – calculated from images acquired on two different generations of Varian EPIDs – are identical within ±1 SD.

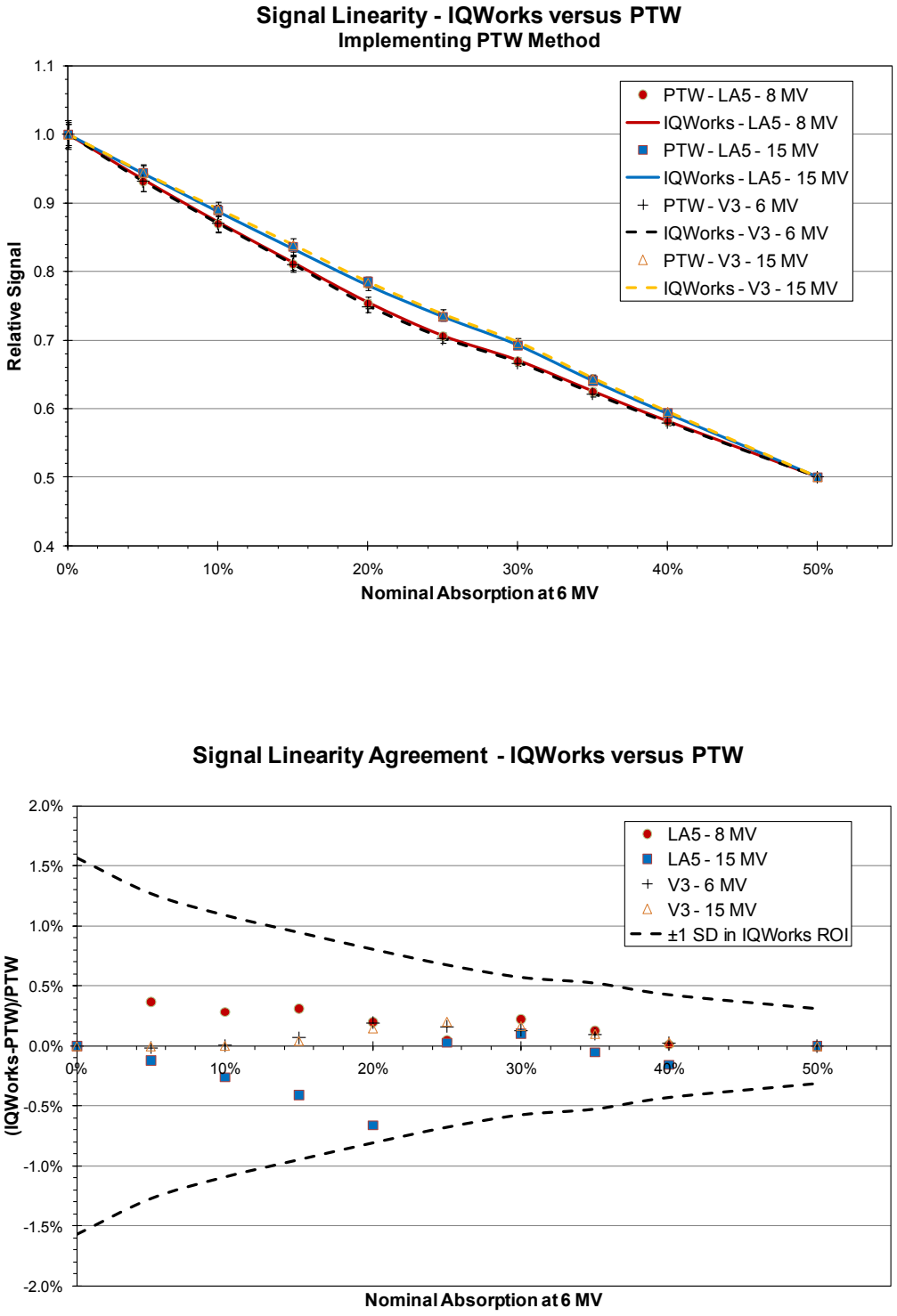


Figure 3.15. – Comparison of IQWorks and PTW epidSoft in calculating signal linearity via the PTW epidSoft method. Four energies across two linear accelerators are considered. Top–Relative signal curves. Bottom–Agreement between IQWorks and PTW epidSoft.

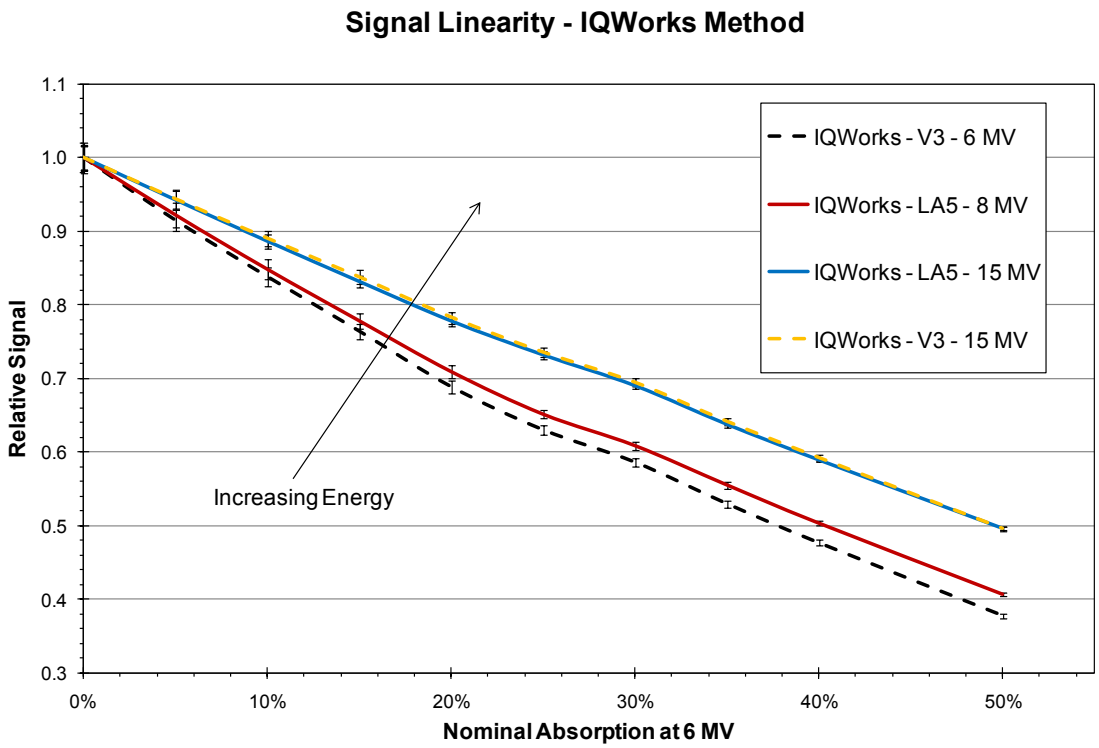


Figure 3.16. – Signal linearity calculated by IQWorks using an alternative method. The same images are considered as in figure 3.15.

3.4.4. Local Signal Linearity: Comparison of IQWorks with PTW epidSoft

Local signal linearity is calculated by epidSoft using a similar methodology to that for general linearity, except that signals from nominal transmissions 10% (block A) to 40% (block D) are mapped to the range 0.6–0.9. i.e.

$$S'_i = 0.3\left(2 + \frac{S_i - S_A}{S_D - S_A}\right) \tag{3.5}$$

A comparison against IQWorks and epidSoft for the 8 MV image acquired in the previous experiment is presented in figure 3.17, in which all IQWorks curves agree with those of epidSoft to within ±1 SD. However, as for signal linearity, this methodology is limited because the first and last points are effectively lost except that the effect is even more pronounced here because there are only 4 points in each curve!

Normalising to unity under block D at corner 1 produces the curves in figure 3.18. It is immediately clear that this alternative analysis yields considerably more information: whilst the curves display the same characteristics across the FOV the relative signal levels increase with proximity to the central axis. This may be a side-effect of using an in-air exposure for flood-field calibration, or may be due to a more penetrating beam spectrum closer the central axis. However, it is an important result which may affect clinical images and is worth tracking over time.

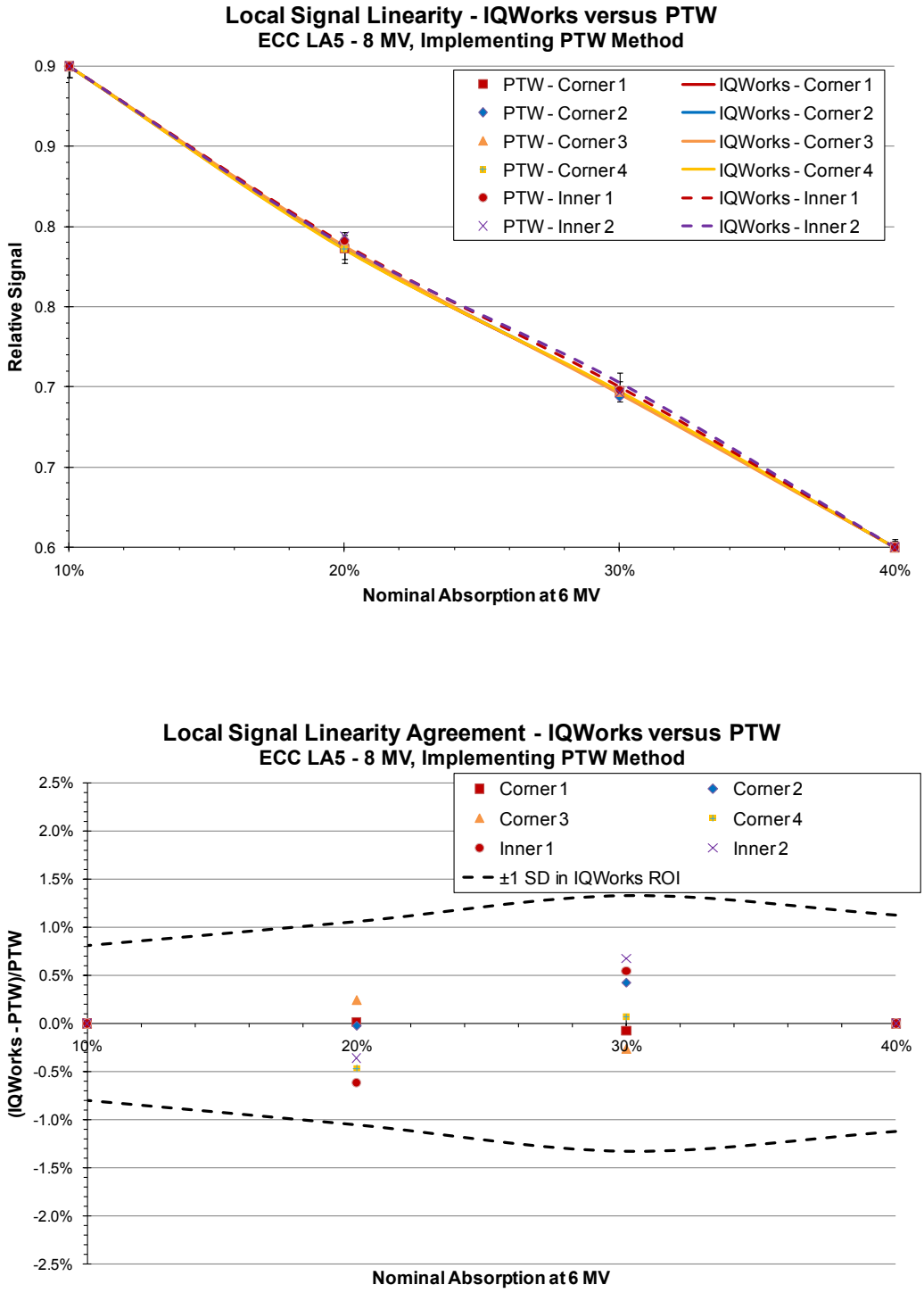


Figure 3.17. – Top–Relative signal curves. Bottom–Agreement between IQWorks and PTW epidSoft.

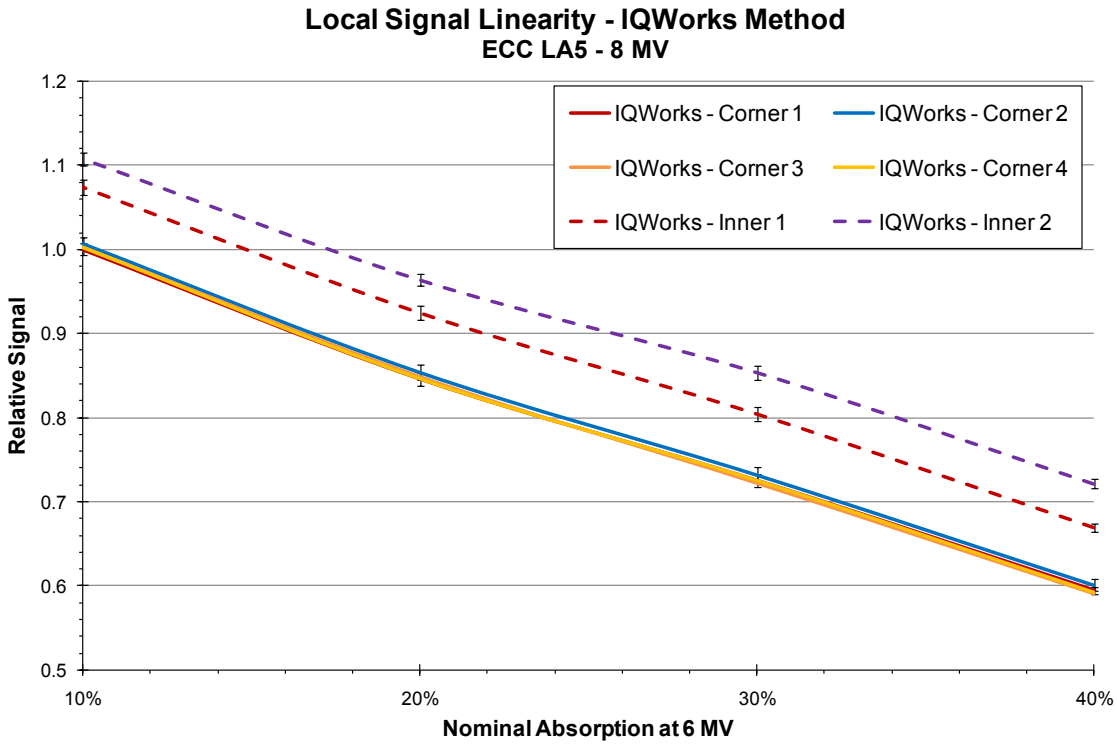


Figure 3.18. – Local signal linearity calculated by IQWorks using an alternative method. The same image was considered as in figure 3.17.

3.4.5. SNR: Comparison of IQWorks with PTW epidSoft

epidSoft evaluates SNR in each of the copper steps S1–S10 using equation 2.3. However, due to both the alignment and image handling issues explored in section 3.4.2 above, it was anticipated that this analysis may be problematic.

An image was acquired using the aS500 detector on linac LA4 in Edinburgh, at 8 MV, 100 MU / min and at a dose of 3 MU (4 frame averages, SQ sync mode). By default, epidSoft produced SNR results which were clearly wrong – SNR systematically improved with reduced signal level, as shown in figure 3.19. It was suspected this was due to poor image handling, and indeed the trend could be replicated in IQWorks through analysing a bitmap image in which the 16 bit image data in the original DICOM file had been remapped to 8 bits. Saving the image as a signed integer matrix then loading this into epidSoft generated a curve more in line with expectations. However, in the same way as for the MTF evaluation it was found that the results were very sensitive to ROI placement. Two IQWorks SNR curves are also plotted in the figure, calculated from two sets of ROIs shifted 1 pixel with respect to each other. These lie either side of the epidSoft curve and follow the same trends, demonstrating agreement with IQWorks within the limits of experimental uncertainty.

At first it may seem surprising that SNR is so sensitive to ROI placement. Although absolute signal level (as the mean pixel value within an ROI) is relatively independent of ROI position, the noise (as the standard deviation in the ROI) is much more variable. This is illustrated for low and high signal ROIs (those under S1 and S10) in figure 3.20, which shows the relative change in signal and noise as the ROIs are moved through a series of locations within the boundaries of the steps. Because the standard deviation is very small (of the the order of 10^1) relative to the signal (of the order of 10^3) any fluctuation in standard deviation can have a profound influence on SNR.

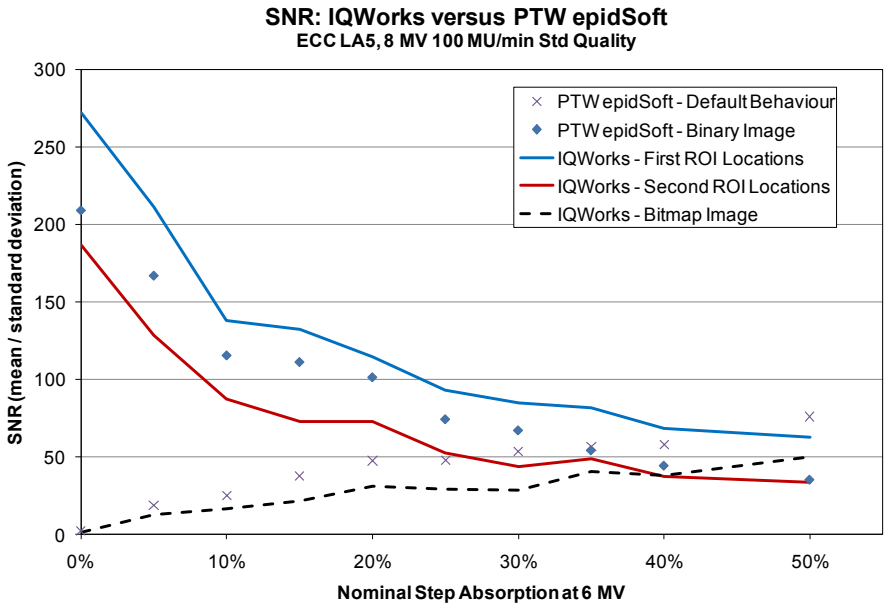


Figure 3.19. – Comparison of Signal to Noise Ratio (SNR) calculated by IQWorks and PTS epidSoft.

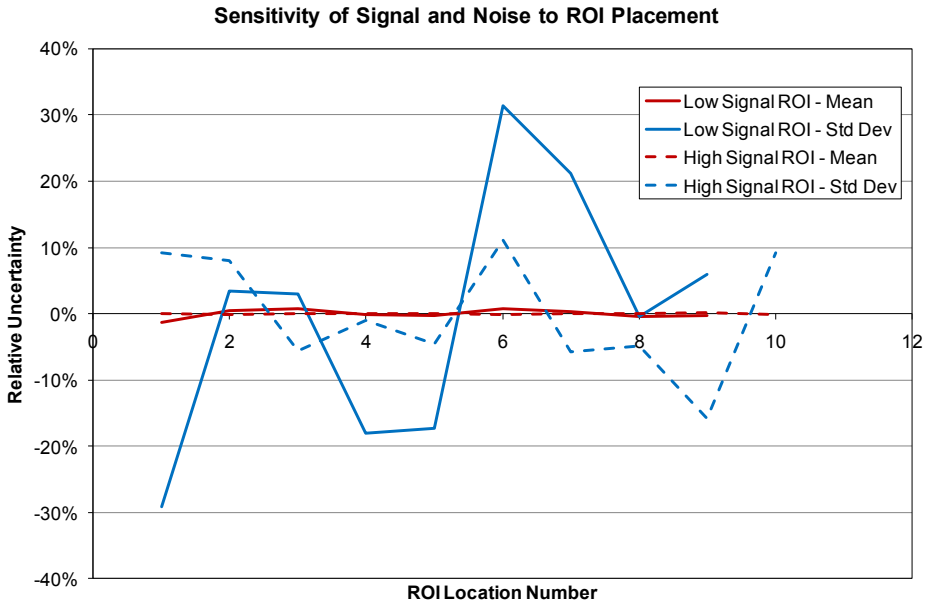


Figure 3.20. – Sensitivity of mean signal and noise (as 1 SD) to region of interest (ROI) placement.

3.4.6. Contrast: Comparison of IQWorks and epidSoft

Contrast in the contrast-detail pattern C1 is assessed by epidSoft using equation 2.8. The signal for a particular hole is taken as the mean pixel value in a circular ROI centred on that hole and the background signal as the mean in an annular ROI surrounding it. Excellent agreement was found between IQWorks and epidSoft when measuring the contrasts of the first 12 holes in the 8 MV image acquired in the previous experiment, as shown in figure 3.21. Better agreement might be expected here than for the SNR and MTF results discussed previously because the pixel values in the region of the aluminium plate containing the contrast-detail pattern are of the order of those in step S8 of the copper wedge, and so are outside of the range affected by the data-handling issues.

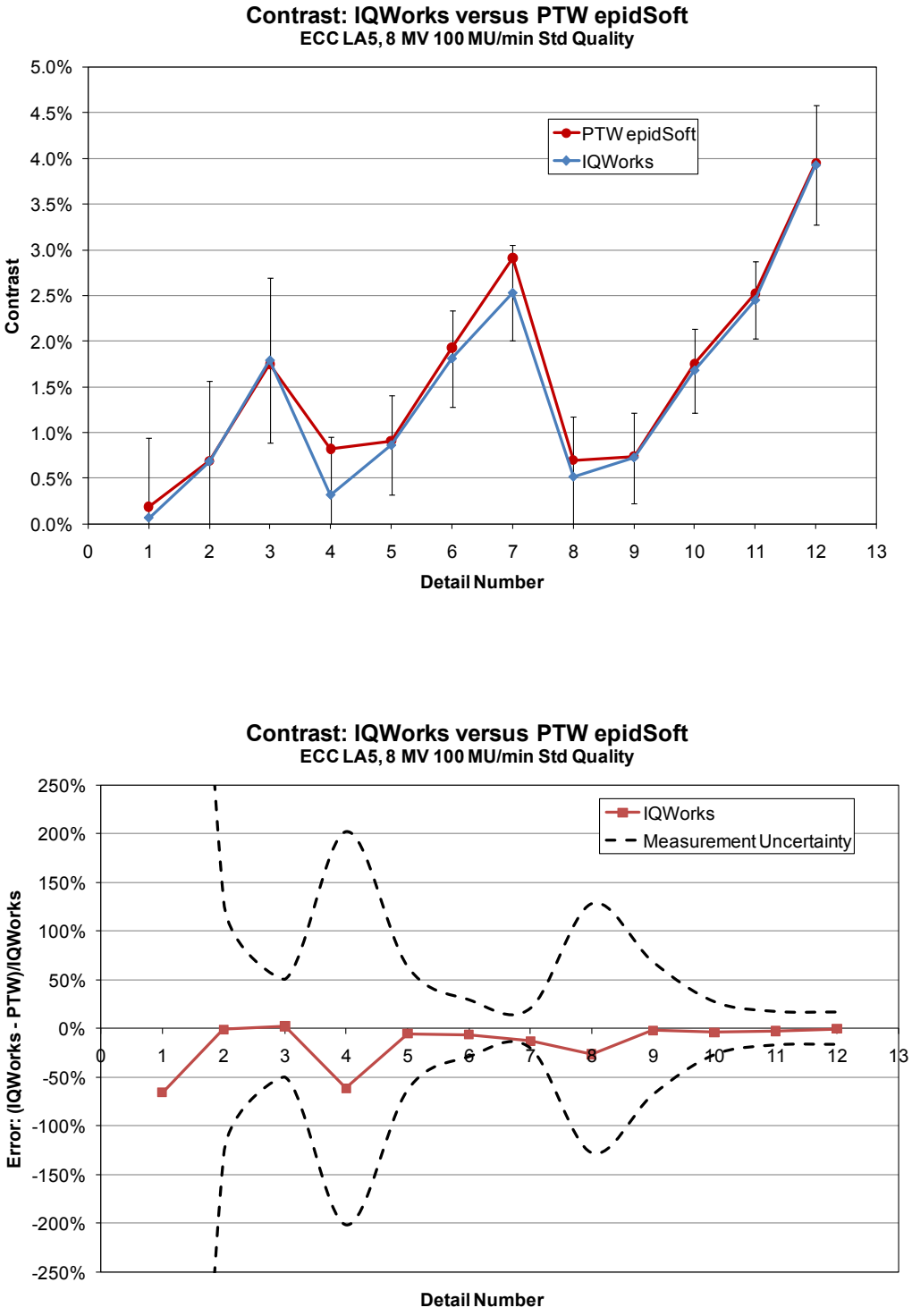


Figure 3.21. – Comparison of IQWorks and PTW epidSoft measurements of contrast in the contrast-detail pattern. Top–Measured contrast for each detail. Error bars indicate measurement uncertainty corresponding to ±1 SD in detail and background ROIs. Bottom–Agreement between IQWorks and PTW epidSoft.

3.4.7. Geometric Linearity

Interestingly, epidSoft does not use the geometry blocks G1–G16 to perform an assessment of geometric linearity even though this is such an important factor in radiotherapy imaging. However, this analysis is relatively straightforward in IQWorks. Following identification of the centres of the blocks by edge detection, the distance to G1 of each block in the vertical series was calculated, and of each block in the horizontal series to G9. Each distance was then divided by the number of blocks between that being considered and the appropriate reference (G1 or G9) to give the distance per block step.

This analysis was undertaken on images acquired using four beams on two linacs: 8 MV (100 MU/min, 4 frame averages, 3 MU) and 15 MV (500 MU/min, 10 frame averages, 8 MU) using the aS500 detector operating under sync mode on LA5 in Edinburgh, and 6 MV (100 MU/min, 2 frame averages, 1 MU) and 15 MV (600 MU/min, 4 frame averages, 2 MU) using the aS1000 detector under rad shot mode on V3 in Oxford. From the results plotted in figure 3.22 it is clear that the inter-step distance falls within a range of 0.5 mm, which is of the order of one pixel, with this becoming tighter over a larger number of steps. Furthermore, there is good agreement between all energies and both linacs.

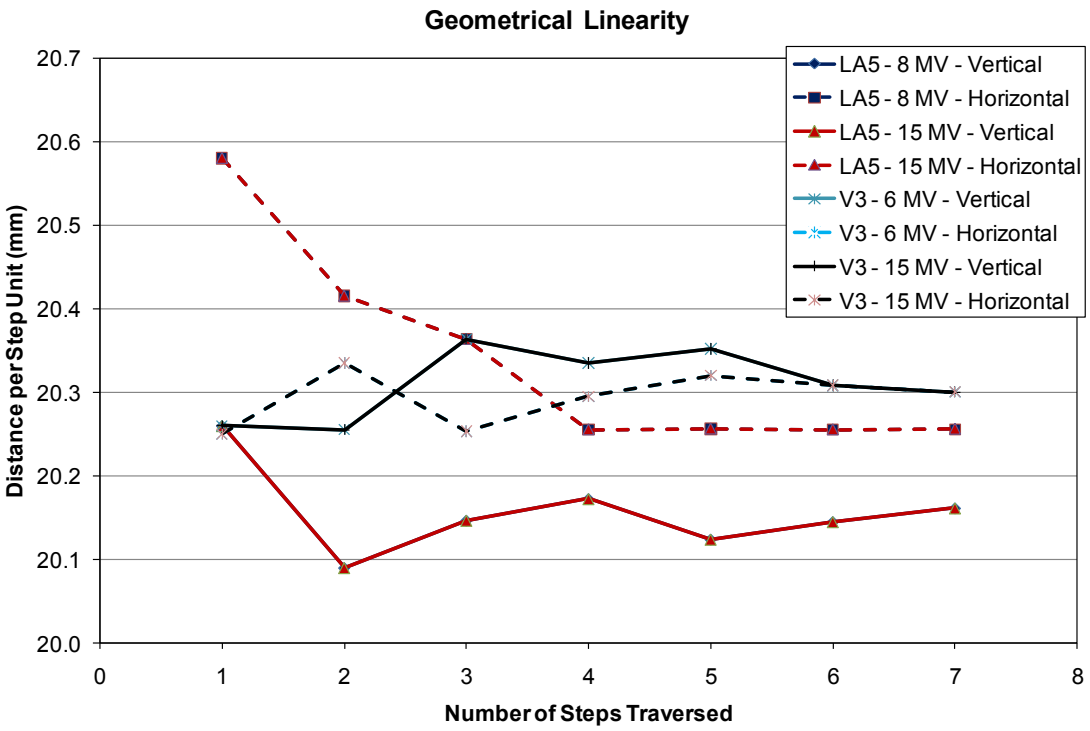


Figure 3.22. – Geometric linearity calculated by IQWorks from images of the PTW EPID QC phantom.

3.5. New 'QEPI1' Phantom

3.5.1. Phantom Introduction

A new phantom for EPI performance evaluation, the 'QEPI1' phantom, was developed as part of this work. Designed specifically to be straightforward and inexpensive to manufacture, the QEPI1 consists of a simple lead square with a square hole at its centre, with the edges of the square angled at approximately 3–4° to the axes of the detector matrix. It is capable of being used to calculated MTF, CNR, SNR and signal and geometric linearity. Furthermore, it can be positioned anywhere in the field of view, at isocentre or on the surface of the detector, and there is no restriction on detector position. A photograph and sample image of the QEPI1 is included in figure 3.23 and schematic diagrams highlighting the features of the phantom used in its analysis are in figure 3.24.

Key features of the phantom include:

- A well-defined geometrical test object which can robustly be found via a range of edge detection algorithms. The Centre of the Extremes (COE) of the phantom edge can be utilised to assess phantom alignment against the central axis position. Although the Centre of Mass of the phantom could also be employed it was found that the COE is less sensitive to edge anomalies due to bad pixels or other artefacts in the image. (See section B.5 for more information on edge detection algorithms and phantom localisation.)

- Geometric linearity in both matrix directions can be assessed by comparing the width and height of the detected phantom edge (as defined in figure 3.24) with the expected values.

- 8 extended edges – 4 round the outside of the phantom and 4 at its centre – distributed across the field of view. The MTF can be calculated from these using the IQWorks 'Edge Line MTF' module. (See section B.17.3.) An advantage of this approach over bar pattern phantoms is that because the MTF analysis is localised to small ROIs, it enables a more precise assessment of MTF at different positions across the field of view. In addition, it is possible to investigate the influence of image or detector artefacts on the local MTF curve by positioning the phantom so that one of the edges aligns with the artefact. With bar pattern phantoms the effect can only be considered on one spatial frequency at a time.

- Regions of interest can be placed anywhere on the lead surface, in the central hole or outside of the square itself. Statistics calculated in these ROIs can be used to assess absolute signal level under different acquisition conditions. Other ROI based analyses can also be performed, such as contrast between the lead and the hole, CNR, SNR, signal gradient and uniformity.

An IQWorks analysis tree was developed to process images of the QEPI1 phantom. This calculates:

- Phantom alignment based on COE of the outer edge.
- Geometric linearity based on the width and height of the extremities of the outer edge.
- MTF curves for all 8 edges, along with the average MTF for the groups of 4 outer and 4 inner edges, and the global average. Standard deviations at each spatial frequency are also determined for these groupings.
- The contrast between the 4 arms of the square and central hole, using equation 2.7 (a simple difference calculation).
- CNR between the 4 arms of the square and the central hole, using equation 2.5 (difference / standard deviation), where the noise level is taken as the standard deviation in the hole ROI only.

Prototype phantoms were constructed which were intended to just fit into the FOV of a Varian EPID when the phantom is positioned at isocentre and the detector at 140 cm SSD. For this scenario, the lead square had an outer dimension of 18 cm, with the inner square being of side 6 cm. However, the actual dimensions may be varied depending upon the portal imaging equipment involved and the details of the particular application. Also, the phantom can either be used in isolation or embedded as a component of a larger test object. For example, in Oxford Cancer Centre the QEPI1 is mounted between PMMA sheets in a larger box which includes alignment markings and other features to allow the unit to be used as a comprehensive daily test device. This enables detailed EPI performance evaluation to be routinely undertaken on a regular basis alongside other key checks. Figure 3.25 illustrates two sample configurations of the QEPI1: the phantom mounted on a jig attached to the head of the linac; and the daily check jig developed in Oxford.

Some flexibility is also afforded in the thickness of the lead sheet, with the ideal being a balance between a sheet which is sufficiently thick that its attenuation at energies used to image the phantom gives good contrast between the material and the central hole, and a sheet which is thin enough that penumbra at the edges does not significantly influence an MTF analysis. Good results for linac beams of energies 6–15 MV were obtained using sheets manufactured from rolled lead of thickness 3.5 mm (BS EN 12588:2006 Code 8 – 'Orange' lead[40]) and 5 mm. However, it is important that the sheet thickness is uniform to within a tight tolerance across its surface. Although BS EN 12588:2006 specifies a tolerance of 5% (approximately 0.2 mm) the phantoms in this work were manufactured to better than 0.1 mm, which is easily achievable through additional rolling by standard machine-shop equipment. Furthermore, it is also crucial that the lead provides uniform attenuation across the surface of the phantom. Experiments were performed using rolled lead and by pouring molten lead into a mould and allowing it to cool. It was found that the rolled lead yielded a more uniform product but that care was still required to avoid voids or other imperfections.

Because MTF curves are determined from the QEPI1 edge spread functions all edges of the lead square should be as sharp and perpendicular to the plane of the square as possible. Generally, the advice when performing such an experiment for diagnostic radiographic imaging is to employ an edge which is as straight and precisely defined as possible – being cut, machined and polished – using a material which is radio-opaque at the beam energies being considered. Tungsten is commonly utilised because of its high density and hardness. Although difficult to machine it yields a very sharp, high quality edge[146, 308]. However, at megavoltage energies it is not possible to manufacture an edge which is sharp across the whole field of view and yet which is even a close approximation of being radio-opaque. Some researchers have attempted to define focused slits or edges using blocks of tungsten, steel or lead, of the order of 10s of centimetres thick[311, 357]. Whilst these succeed in being radio-opaque this approach is of limited value because the slit or edge is only defined along a single line across the detector and the experimental arrangements are too inconvenient and cumbersome for routine performance evaluation. However, there is a growing consensus in the diagnostic imaging community that for all but absolute reference performance evaluations, it may be acceptable to utilise non-ideal attenuating edges: those which are not perfectly polished or radio-opaque. Various researchers have experimented with cruder copper, steel

and lead edges and have yielded results in good agreement with evaluations performed under more stringent experimental conditions[78, 179, 234, 257, 271]. It was therefore hoped that the machined, perpendicular edges of the lead square would be sufficient for this application. In particular, focusing the edges back to the radiation source was not thought to be of great importance as long as the lead was sufficiently thin, with the 3.5 mm and 5 mm thick prototypes being approximately one third and one quarter the thickness of the equivalent bar pattern details in the QC-3V and PTW EPID QC phantoms respectively, neither of which are accurately focused back to the radiation source.

As will be described below, good results were achieved using the simply machined lead phantom, and there was good agreement between MTFs calculated by the QEPI1 phantom and a tungsten edge when imaged at kilovoltage energies, as described in chapter 5. Furthermore, the initial QEPI1 prototype has continued to produce reproducible results over a number of years even though it has been roughly handled and the edges occasionally misshapened. The central hole can therefore be cut by punching out a portion of lead and finishing the edges by machining.

Although edge angles must be accurately known in order to determine edge spread functions, because IQWorks automatically localises the line describing each edge it is not important that the lead sheet be precisely angled during construction of the phantom. As long as the angle is approximately 3–4° then there is sufficient oversampling to yield good MTF results whilst maintaining a large enough sampling interval to minimise the impact of noise on what is an inherently low contrast system. Furthermore, automatic edge characterisation by IQWorks means that the analysis is relatively tolerant of small angular phantom misalignments.

Overall, the flexibility of the QEPI1 phantom, together with the whole range of evaluations which can be performed, allows it to be a valuable tool in a wide range of quality assurance and optimisation experiments. Being readily accessible to all radiotherapy departments it is hoped the QEPI1 phantom will encourage the adoption of a consistent and objective approach to EPI image quality assessment.

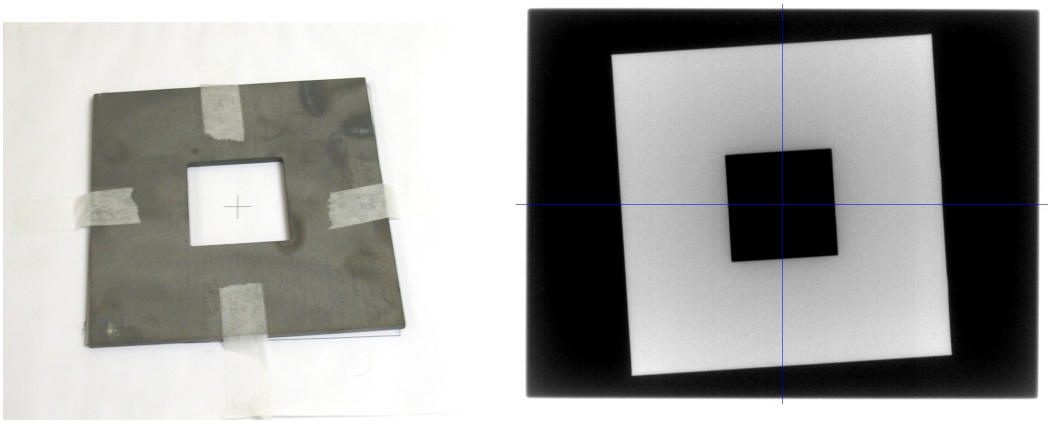


Figure 3.23. – The new 'QEPI1' phantom. Left–Photograph of lead square at the heart of the phantom. Right–6 MV EPI of the phantom.

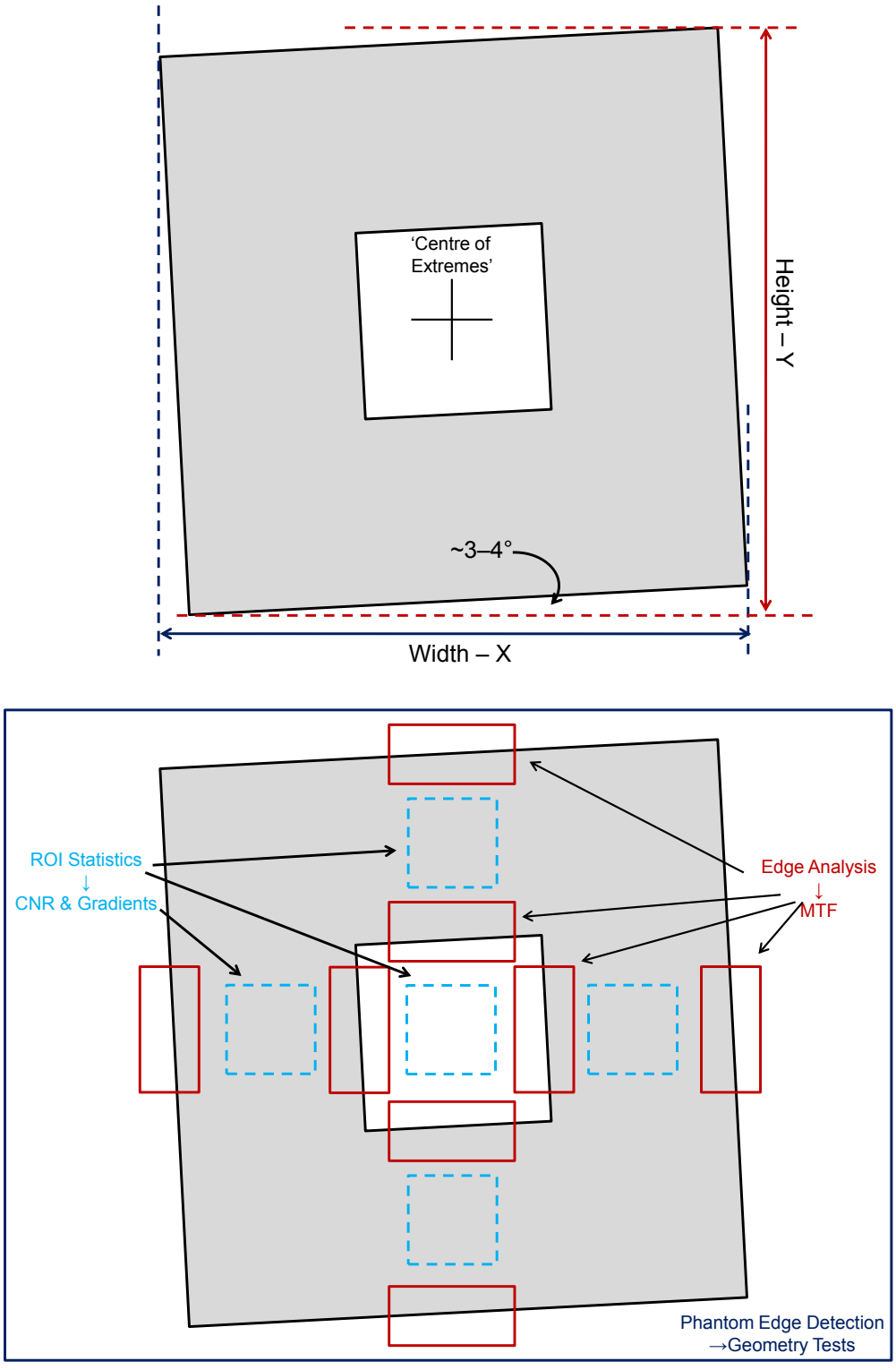


Figure 3.24. – Schematic diagrams of the QEPI1 phantom. Top–Illustration of phantom angulation and definition of 'Centre of Extremes', 'width' and 'height' for geometrical assessments. Bottom–Regions of interest for analysis.

Figure 3.25. – Options for aligning the QEPI1 phantom at isocentre. Left–The phantom is mounted on an assembly attached to the linac gantry. Right–The phantom is embedded in a PMMA jig which is positioned on the treatment couch.

3.5.2. Development of QEPI1 Jigs

Six daily check jigs were constructed, one for each of the six linear accelerators in the new Oxford Cancer Centre and each incorporating a QEPI1 test object made from Code 8 (3.5 mm) lead. In order for the jigs to be used interchangeably, or for it to be viable to compare measurements made with one jig against those performed using another, an intercomparison was undertaken by imaging all jigs under controlled conditions using all clinically available imaging modes on two linacs.

Before performing detailed image quality measurements it was necessary to verify that the lead of QEPI1 squares did not contain any imperfections which might affect the analyses. Variance maps were generated of the QEPI1 images acquired at 6 MV, 100 MU/min in High Quality Rad Shot mode (the average of 4 frames), at a dose of 2 MU. This imaging protocol was chosen because it was the highest contrast, lowest noise option available, as discussed in section 3.2. Nearest neighbour pixels (i.e. ±1 either side of the reference pixel) were considered when calculating the variance maps and variances above 100 were truncated to prevent high values at the edges of the squares suppressing low-level structure. (See section B.16 for details of the variance map calculation algorithm.)

All the variance maps are shown in figure 3.26 and it is clear from visual inspection that there are no clusters of structure on any of the squares. A number of vertical lines are are visible but these were attributed to columns of defective pixels because they were still present in images taken when the jigs were removed. From these results there is confidence that each jig is sufficiently uniform to be employed in ROI based image quality analyses.

Images of all jigs were acquired using the aS1000 detector on a single linac (V3) using all acquisition modes which might be encountered clinically. Due to its inherent advantages over the older Sync mode (i.e. improved image quality at a lower dose, as described in section 3.2) it was decided to employ the new IAS3 Rad Shot mode for all patient set-up verification imaging, so this is the only panel readout mode considered here. In summary, six acquisition modes were utilised:

- 6 MV, 100 MU/min, Standard Quality (2 frames, at a dose of 1 MU)*
- 6 MV, 100 MU/min, High Quality (4 frames, at a dose of 2 MU)
- 6 MV, 600 MU/min, Standard Quality (2 frames, at a dose of 1 MU)

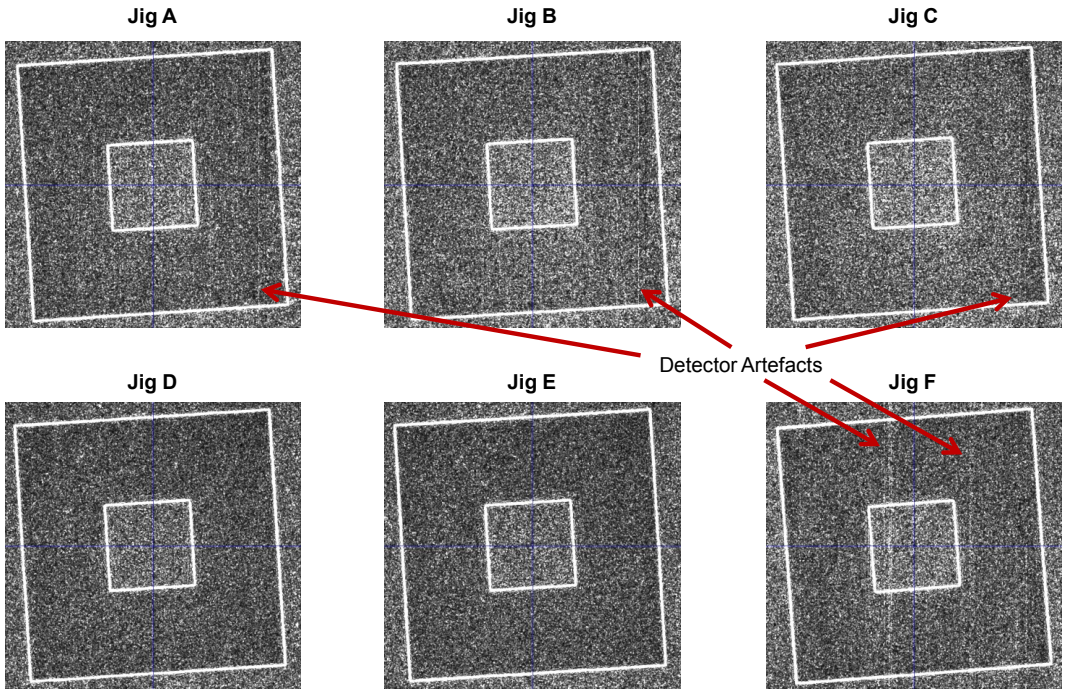


Figure 3.26. – IQWorks variance maps for each of 6 QEPI1 jigs. Artefacts due to columns of bad pixels in the detector are highlighted.

- 6 MV, 600 MU/min, High Quality (4 frames, at a dose of 2 MU)*
- 15 MV, 600 MU/min, Standard Quality (2 frames, at a dose of 1 MU)
- 15 MV, 600 MU/min, High Quality (4 frames, at a dose of 2 MU)*

Those indicated by an asterisk (*) are the ones used clinically by default, although all of these are available clinically and there is scope to switch between them if the application warrants doing so.

Charts of contrast and contrast-to-noise ratio (CNR) for each imaging mode, averaged across all jigs, are presented in figure 3.27. As expected, there is no significant difference in absolute contrast between the 2 and 4 frame modes, for any given energy and dose-rate. However, an interesting result is that there is a significant boost in contrast when acquiring 6 MV images at the higher dose-rate. This is surprising because the beam spectra and thus attenuation conditions are the same at both dose-rates, so that subject contrast should be the same. Furthermore, by design the IAS3 Rad Shot mode causes the detector to be read out at a constant frame-rate corresponding to a nominal dose per frame, as described in section 3.2. Because the whole field of view is subject to the same beam conditions there should be no influence from beam stabilisation effects at the higher dose-rate, as may have been encountered with older, IAS2 based

detectors[343]. A marked reduction in contrast is observed between 6 and 15 MV, which is as expected because the higher energy beam is more penetrating.

There is a clear boost in CNR between images acquired with 2 and 4 frame averages at all energies and dose-rates. Quantitatively, due to Poisson statistics one would expect CNR to increase by a factor of $\sqrt{2} = 1.4$ when the dose is doubled. In figure 3.28 the ratio of CNR between 4 and 2 frame averages is charted relative to this value and it is clear that the measured boost agrees with expectation, to within the limits of experimental uncertainty. Interestingly, there is no significant difference in CNR between the low and high dose-rates when imaging at 6 MV, despite the slight difference in measured in contrast.

As described in section 3.2, the technique routinely employed for dedicated imaging fields in Oxford is a 6 MV, 100 MU/min beam and acquiring 2 frame averages at a total dose of 1 MU. However, it is clear that averaging 4 frames offers a significant boost in CNR and this has been made available to radiographers as a higher-dose option if the clinical situation merits it.

Interesting results were obtained for the MTF measurements. As expected, there was a significant deterioration in MTF between 6 and 15 MV, with two distinct curves being clearly visible, as shown in figure 3.29. For the groups of inner or outer edges, all jigs produced average f_{50} results which agreed to 2 significant figures. However, when examining the average MTF curves for the inner and outer edges it was apparent these could be resolved into two separate curves lying either side of the global mean, with the curve for the outer edges indicating slightly better performance than that for the inner edges. This finding was consistent at both 6 and 15 MV, although it was marginally more pronounced at 6 MV.

At first, the variation in MTF across the field of view was thought to be due to the penumbra of the edges caused by the 3.5 mm thickness of the phantom, with these being magnified due to the phantom being aligned at isocentre. However, similar results were obtained with the phantom positioned on the surface of the detector, when the effect of penumbra should be minimal. Furthermore, and somewhat paradoxically, the difference between the curves was actually found to be less for a 5 mm thick prototype QEPI1 phantom where one would expect the penumbra to be worse. Next, it was suspected that manufacturing differences between the Oxford jigs may be having an impact, but the same trends were observed for each individual jig, as illustrated in figure 3.29.

That this is a real effect can be demonstrated by considering the edges spread functions for individual edges. ESFs for single outer and inner edges of one of

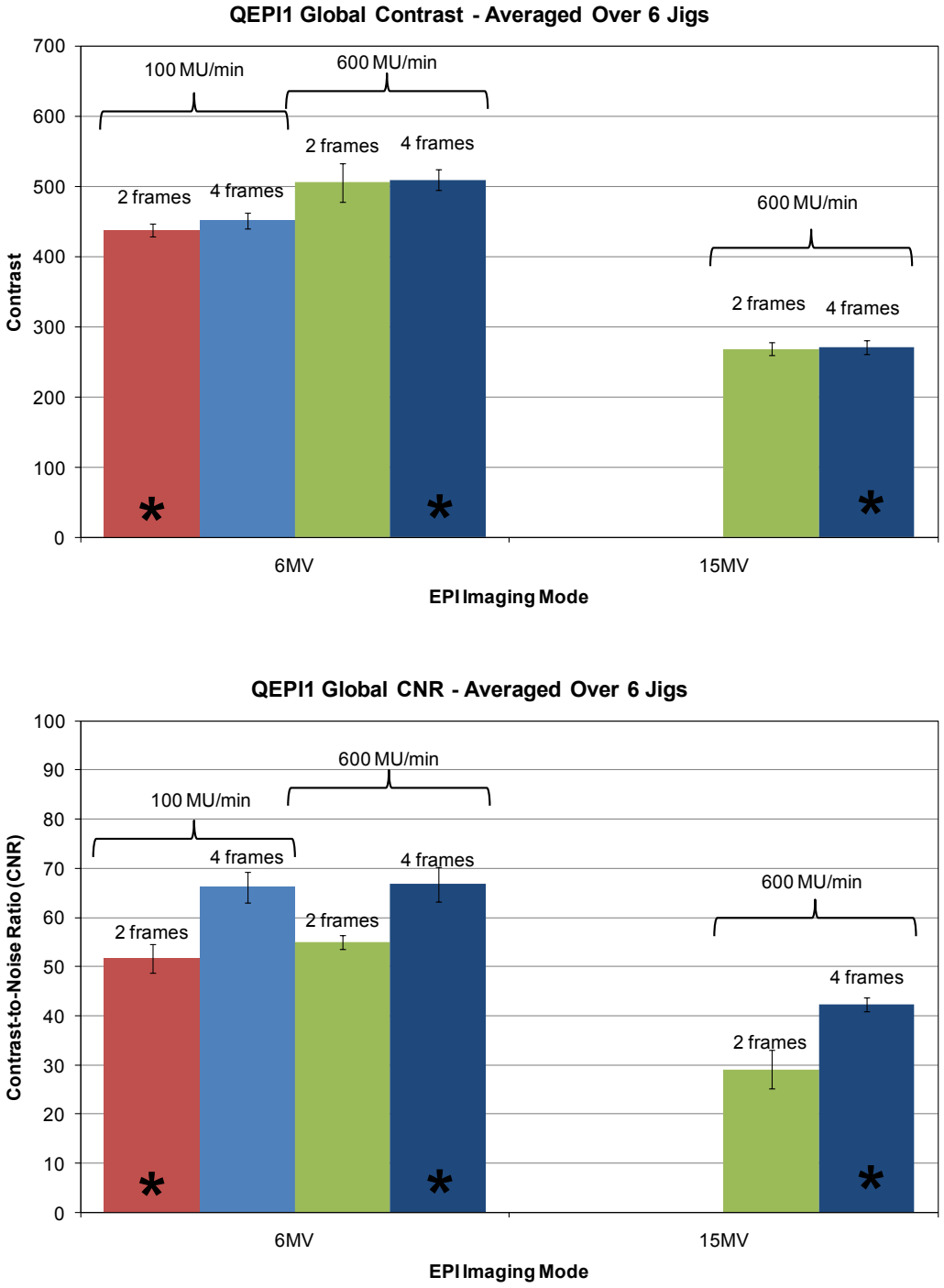


Figure 3.27. – Top–Contrast and Bottom–Contrast to Noise Ratio (CNR) of different Oxford EPI modes, assessed using the QEPI1 phantom and IQWorks. Error bars indicate the 95% confidence interval in the mean value across all jigs. * denotes modes which are routinely used clinically.

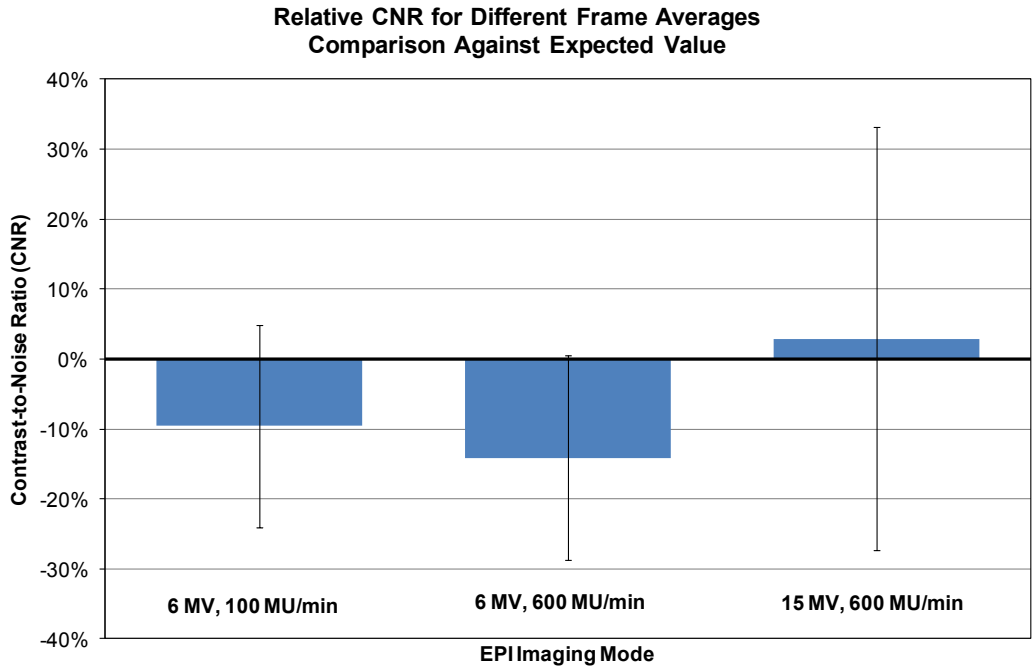


Figure 3.28. – Comparison of the ratio between the contrast-to-noise ratio (CNR) results measured using the QEPI1 jigs for 4 and 2 frame averages against the expected improvement factor of 1.4. Error bars indicate the 95% confidence interval.

the jigs are plotted in figure 3.30. In the plots the curves have been normalised to account for changes in signal level across the field of view or with energy. At both energies the ESF of the outer edge appears slightly sharper than that of the inner edge, consistent with MTF curves from outer edges exhibiting better performance.

A possible explanation for this phenomenon is that the radiation beam spectrum becomes less penetrating with distance from the central-axis[111, 361]. For any given thickness of lead square, the outer edges will therefore attenuate the beam more than the inner, resulting in sharper ESFs and better MTFs. This is also consistent with the thicker QEPI1 square exhibiting less of a difference across the field of view. Although, strictly speaking, irradiation conditions should be such that the edge is completely opaque when calculating MTF this is not achievable at megavoltage energies, with any spatial resolution phantom requiring a pragmatic balance of compromises. It is reassuring that the QEPI1 phantom is sensitive enough to highlight changes in measured MTF due to a characteristic of the megavoltage X-ray beam which might also influence image quality in real patient exposures.

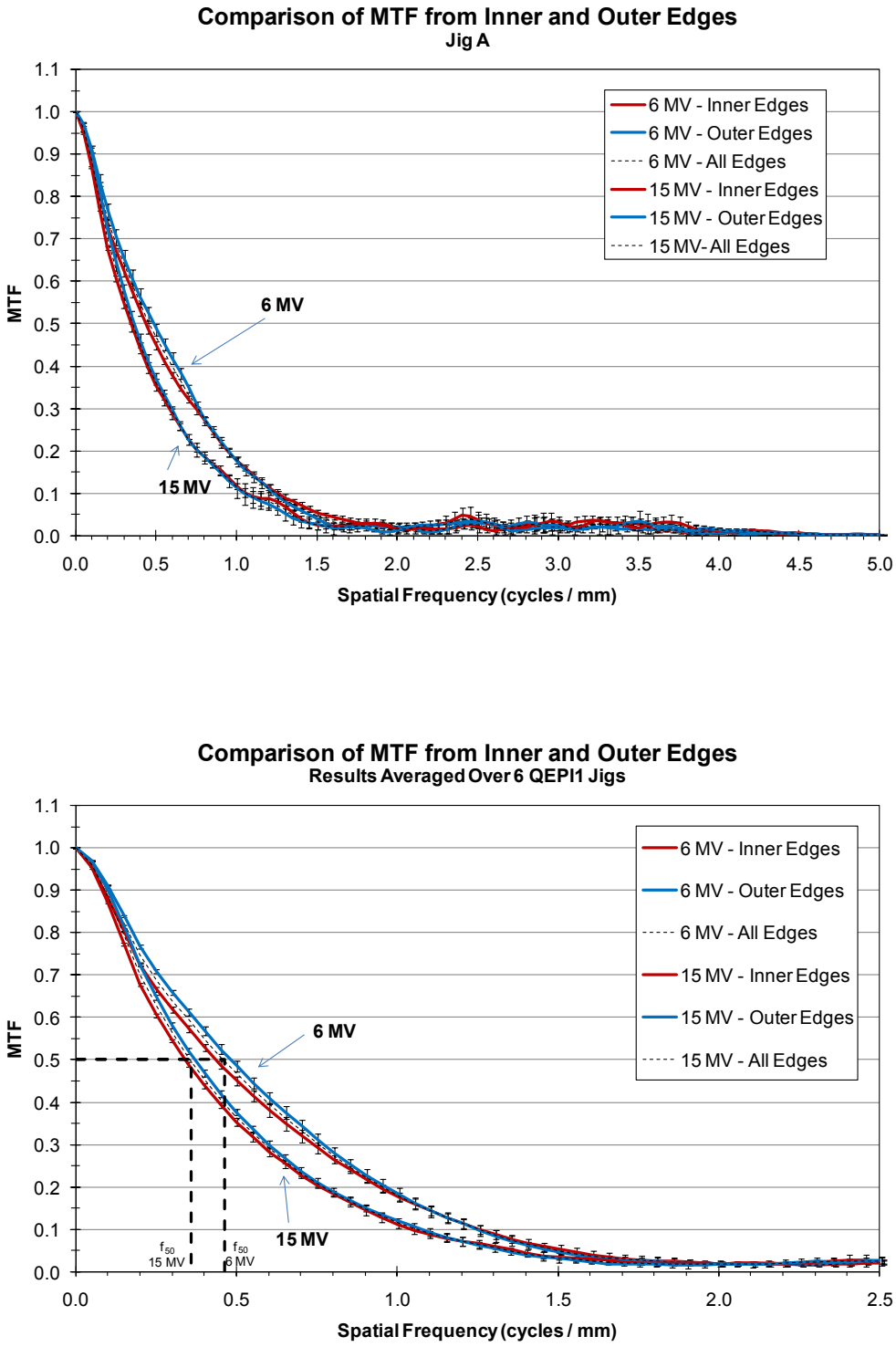
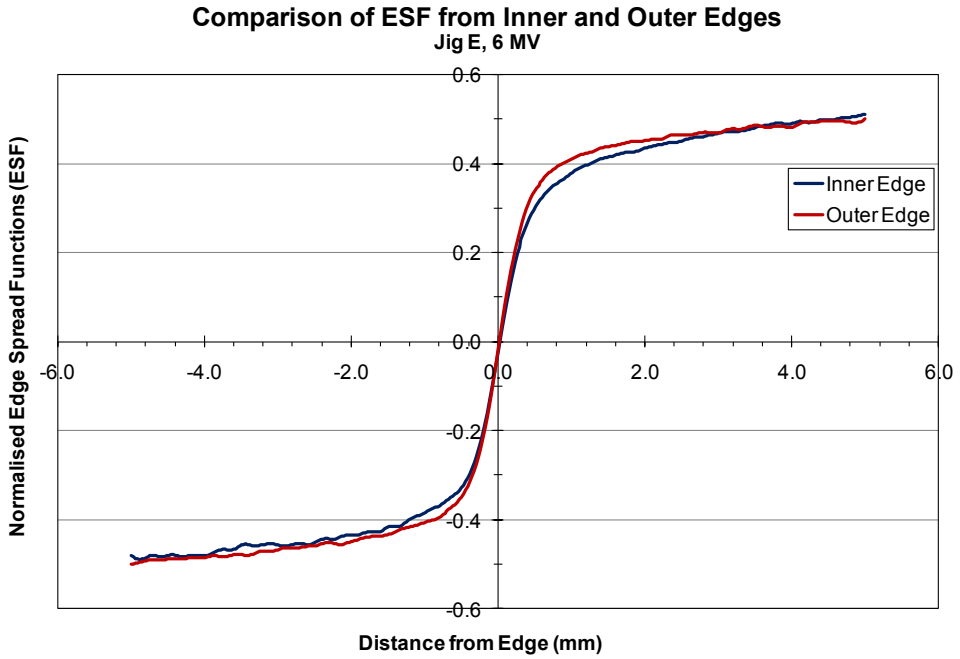


Figure 3.29. – MTF measured across the field of view for 6 and 15 MV imaging modes using the QEPI1 phantom and IQWorks. Top–Results for Jig A only. Bottom–Results averaged across all jigs.



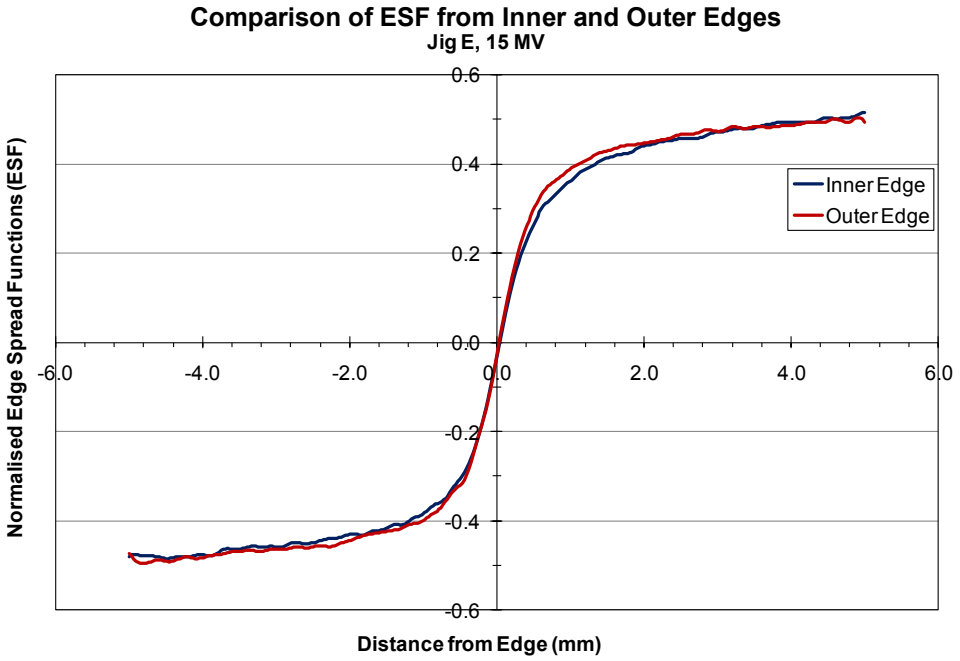


Figure 3.30. – Edge Spread Function across the field of view, measured using the QEPI1 phantom (Jig E) and IQWorks. Top–Results for 6 MV. Bottom–Results 15 MV.

3.6. Comparison of MTF Results from Different Phantoms

In the preceding sections the ability of IQWorks to analyse images of three different EPI phantoms was clearly demonstrated. However, in trying to mimic the commercial software provided with the QC-3V and PTW phantoms the analysis algorithm employed was slightly modified each time, thus preventing direct comparisons between image quality metrics calculated using the various phantoms. Images were acquired of each phantom using the same acquisition settings: a 15 MV, 500 MU/min beam, taking 10 frame averages at a total dose of 8 MU/image. This experiment was performed on linac LA5 in Edinburgh and was the only occasion when all three phantoms were available together.

Raw MTF curves calculated by the dedicated software packages for each phantom are displayed in figure 3.31. Aside from differences in the numerical implementations of the MTF calculation algorithm it is evident from the figures that comparisons are also hindered by the MTF curves being normalised to a different spatial frequency for each phantom – that of the lowest frequency bar pattern in the QC-3V and PTW phantoms (0.1 cycles / mm and 0.125 cycles / mm respectively), and zero-frequency for the QEPI1 phantom.

One of the strengths of IQWorks is the potential to develop analysis trees which yield metrics which are comparable between phantoms. The same images were processed again, this time all in IQWorks using compatible algorithms and with the MTF curves all being normalised to unity at 0.25 cycles / mm. This spatial-frequency was chosen because it is that of one of the bar patterns in both the QC-3V and PTW phantoms, and is one of the points generated by IQWorks for the QEPI1 MTF curve. The experiment was also repeated at 8 MV. From the comparative plots in figure 3.32 it is clear that at both energies the curves from each phantom now agree within the limits of experimental uncertainty (taken as ±1 SD in the MTF values calculated at a particular spatial frequency for all 8 edges of the QEPI1 phantom). This is an important result because it demonstrates that, when using compatible analysis trees in IQWorks, any of the phantoms can be used interchangeably to calculate MTF. However, the normalisation approach is non-standard and would still prevent comparisons against results reported for diagnostic equipment, where the MTF is always normalised to unity at zero frequency.

A further analysis was performed in IQWorks, this time utilising the approach suggested by Droege and Morin[79, 80] to determine the relative modulation

at zero frequency, namely placing ROIs on patches of materials with the same attenuations as those of the bars and gaps in the bar patterns. This was straightforward for the QC-3V phantom (U2 and U6 in figure 3.1) but required some trial and error for the PTW phantom because the height of the brass bars is not documented. For this phantom the best results were obtained when using S3 in figure 3.7 and an additional ROI on the base material. The MTF curves calculated by this approach for each energy are again in agreement within the limits of experimental uncertainty and are plotted in figure 3.33. These curves are now directly comparable against results presented in the literature for diagnostic imaging equipment, demonstrating the universality and value of the IQWorks framework.

Due to the different attenuation characteristics of the materials in the various phantoms it is not possible to calculate SNR or other metrics based on absolute signal values which are compatible between the phantoms. However, a benefit of the QEPI1 being inexpensive and straightforward to manufacture is that it can readily be adopted as a baseline device for undertaking comparative measurements.

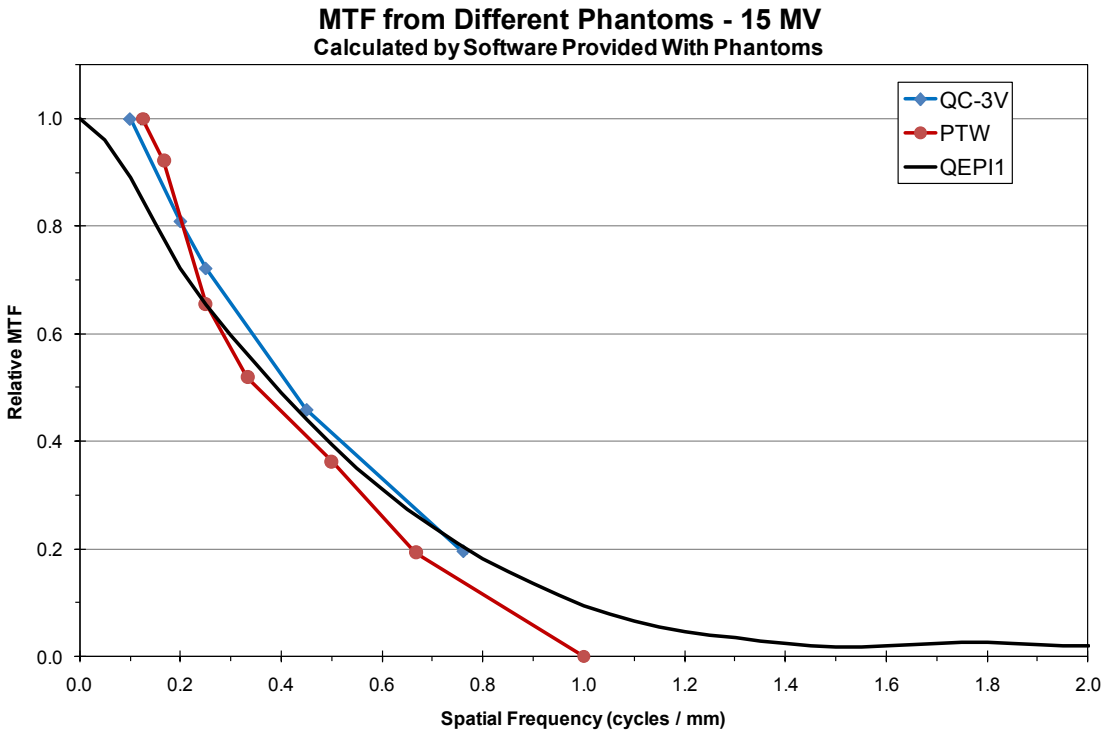


Figure 3.31. – Comparison of modulation transfer function (MTF) for each phantom, calculated using software provided with the phantoms. Images were acquired at 15 MV.

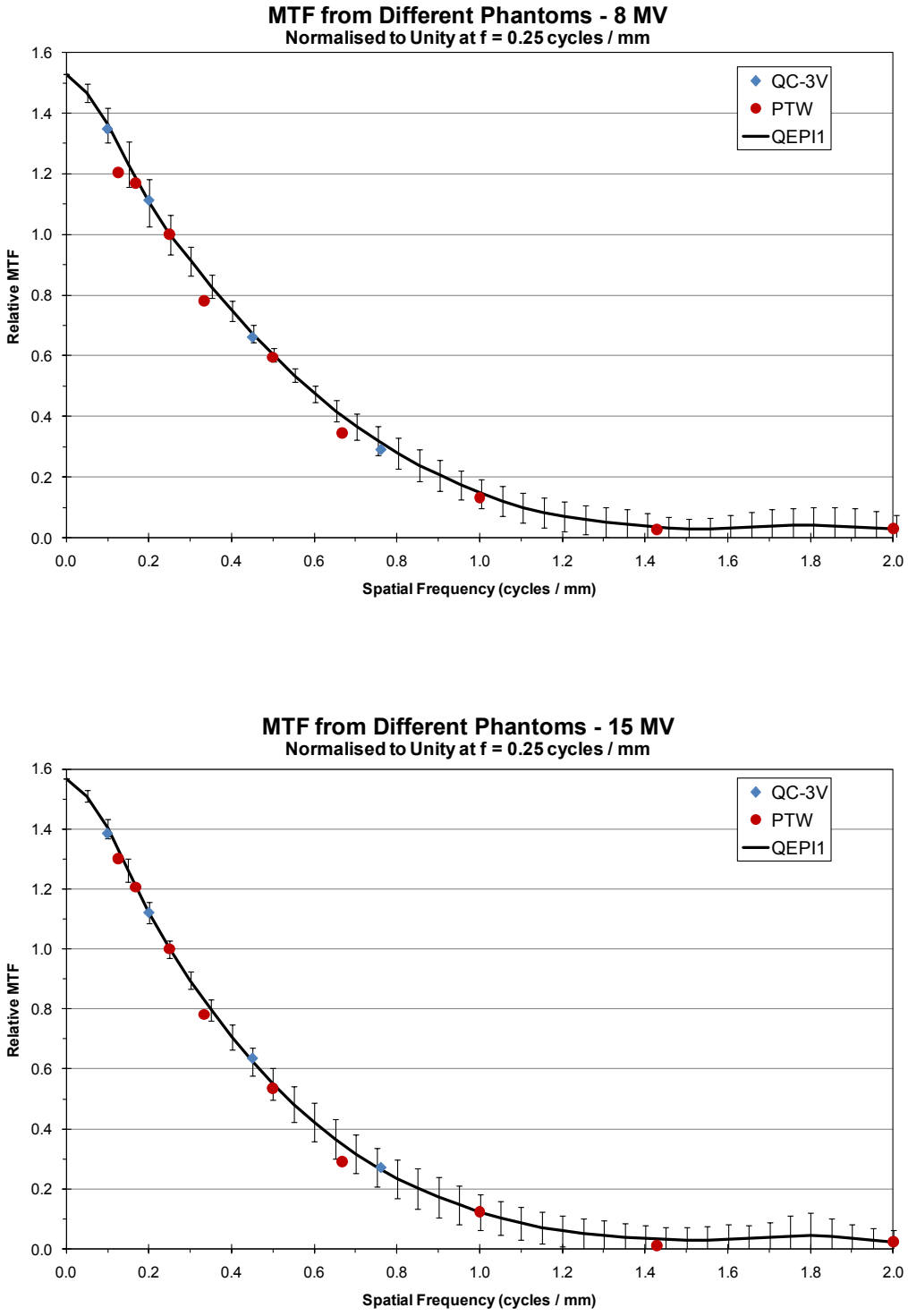


Figure 3.32. – Modulation transfer functions (MTFs) of all phantoms, calculated by IQWorks and normalised to unity at 0.25 cycles/mm. Top–8 MV, Bottom–15 MV.

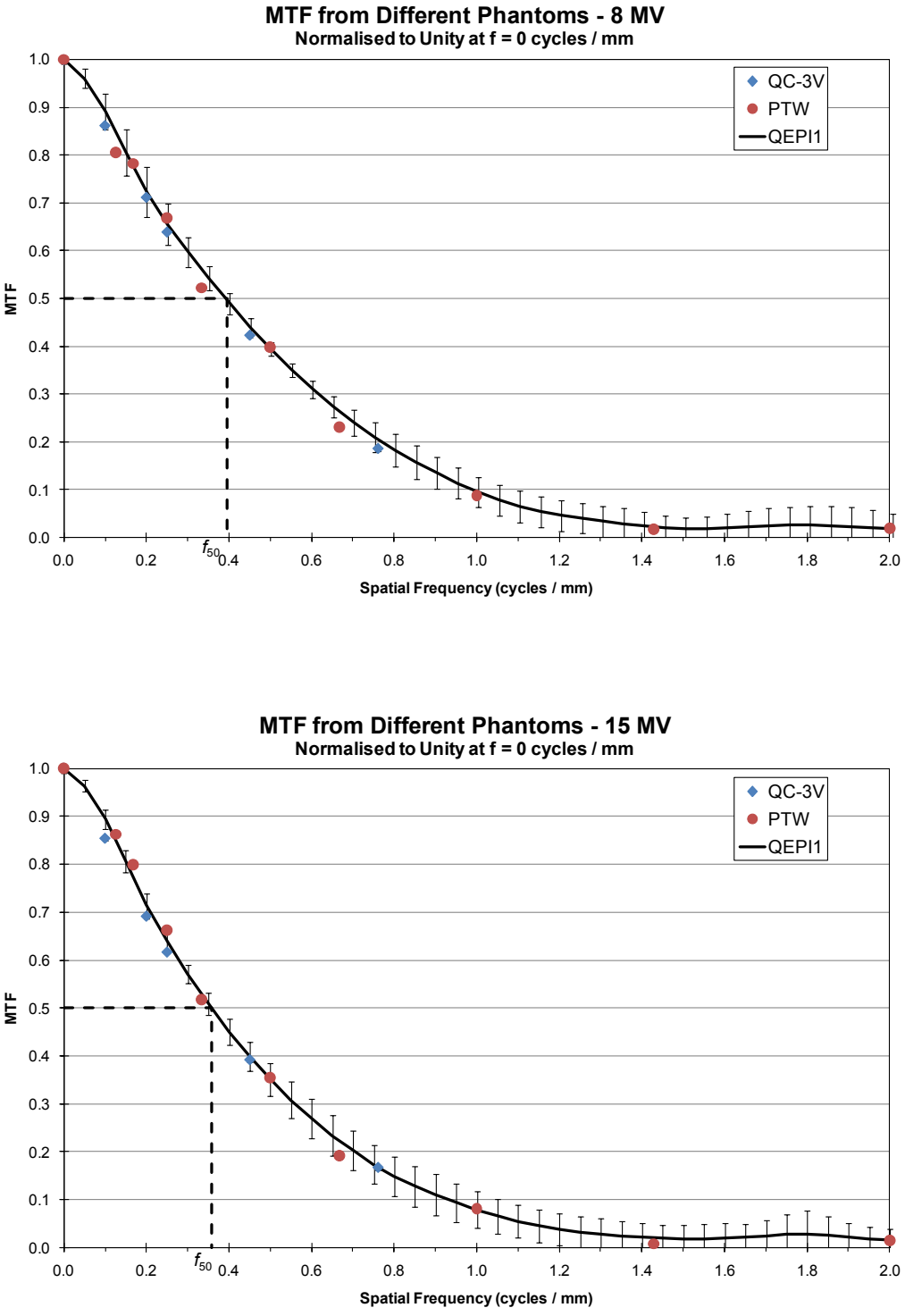


Figure 3.33. – Modulation transfer functions (MTFs) of all phantoms, calculated by IQWorks and normalised to unity at 0 cycles/mm. Top–8 MV, Bottom–15 MV.

3.7. Evaluation of aS1000 Licence Mode

All six linear accelerators in Oxford are equipped with Varian aS1000 EPIDs. As discussed in section 3.2, it was discovered during equipment commissioning that these appear to be the same as the cheaper aS500-II model detectors but with an additional software licence that enables 'full-resolution' readout of the 1024×768 pixel matrix. When this licence is not applied the pixel signals are binned 2×2, giving an image matrix with effective dimensions of 512×384 over the same 40 cm $\times$ 30 cm panel area.

Although Varian market their aS1000 option as being the most appropriate for all clinical applications due to its intrinsically higher spatial resolution, it was suspected that this may be achieved at the expense of CNR. As a result of the smaller pixel size fewer photons can contribute to the signal in any pixel at a particular dose level and it was expected that the relative noise level in the image must be higher. According to Poisson statistics, reducing the pixel size (and hence dose per pixel) by a factor of 4 should result in an increase in relative noise level by a factor of 2, and hence a reduction in CNR by half.

Experiments were performed using the QEPI1 and IQWorks to fully characterise the performance of the EPID with and without the aS1000 licence present. These would enable consideration of whether it might ever be desirable to run in the aS500-II mode and assess the benefit gained by applying the aS1000 licence. As in the previous section, only the newer Rad Shot readout mode was considered because it was not envisaged ever running the detector under the older Sync mode.

Before the licence evaluation was performed the detector was fully calibrated by applying new dark and flood-field corrections. However, it was of interest to identify the influence of a routine calibration on fundamental image quality. Images of the QEPI1 phantom were therefore acquired before and after calibration using all clinical imaging modes. Surprisingly, it was discovered that all QEPI1 metrics were similar, within the limits of experimental uncertainty, before and after calibration. In particular, the MTF curves for each energy were effectively identical, as illustrated in figure 3.34. This was of concern because it was known from experience that calibrating an EPID improved perceived image quality by images appearing less noisy, yet this was undetectable in the standard QEPI1 analysis. As expected, the measured MTF curves at 15 MV were significantly poorer than those at 6 MV. This is due both to increased scatter in the detector and because the brighter flashes of light from the higher energy photons which interact in the phosphor diverge over a larger area before

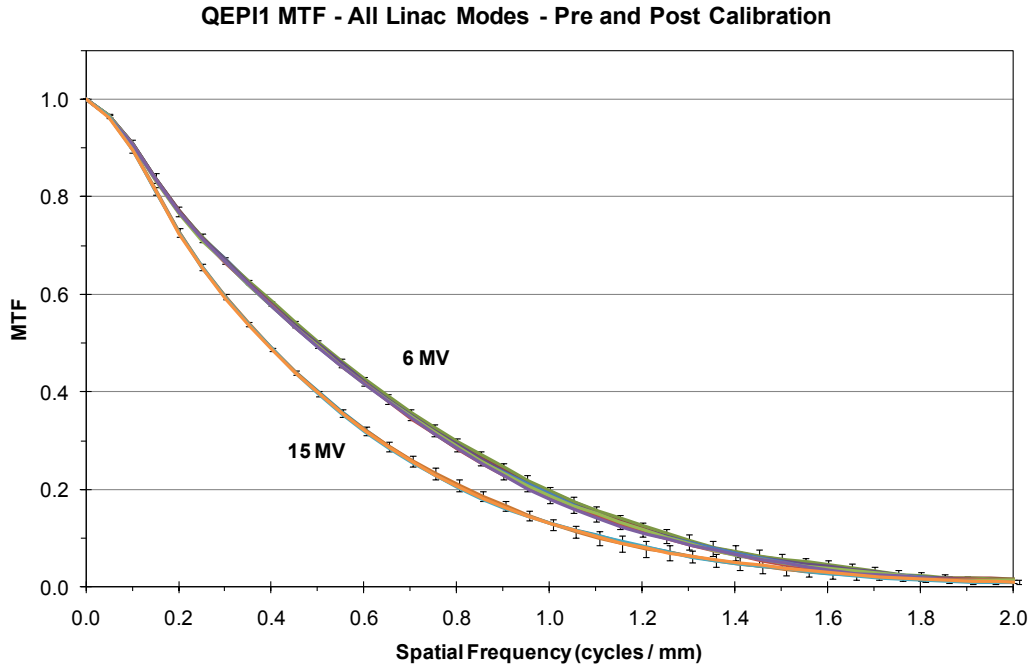


Figure 3.34. – Modulation transfer function (MTF) of a Varian aS1000 EPID for 6 and 15 MV imaging modes, assessed using the QEPI1 phantom and IQWorks before and after comprehensive dark and flood-field calibrations. Curves are the average MTF measured over the field of view and the error bars indicate ±1 SD of the individual MTF values.

interacting with the sensitive amorphous silicon layer, thus adding a blurring effect to the image[239, 249, 356].

A detailed NNPS evaluation was also performed on uniformly exposed images taken before and after calibration using the IQWorks 'NPS Auto ROI' module described in section B.18.4. The module was configured to fill 97.5% of the FOV with half-overlapping ROIs of dimensions 128 × 128. An interesting result from the curves plotted for all imaging modes in figure 3.35, and focusing in on a single mode for clarity in figure 3.36, is that whereas EPID calibration yields only a marginal, if any, improvement in stochastic noise there is a marked reduction in the fixed-pattern noise. This is an important finding because it is generally believed that image quality in electronic portal imaging is limited by stochastic noise[123].

Following a comprehensive calibration, and performance of baseline QEPI1 and NNPS analyses, the aS1000 licence was disabled and a further series of images acquired.

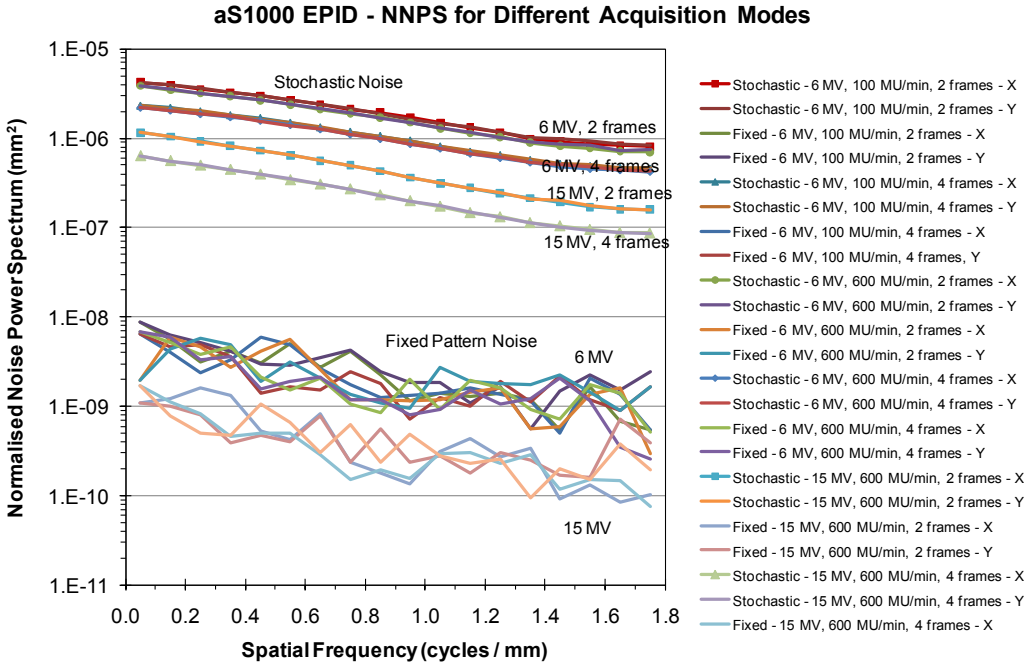


Figure 3.35. – Comparison of stochastic and fixed-pattern components of the normalised noise power spectrum (NNPS) for all 6 and 15 MV 'RadShot' imaging modes of the Varian aS1000 EPID.

MTF curves with and without the licence enabled are presented in figure 3.37, which shows that when the aS1000 licence is present the MTF curves exhibit a significant improvement. This is particularly the case over the range 0.4–1.4 cycles / mm, where the MTF is boosted by a factor of between 50% and 100%. It is thought this might be advantageous for precise image guidance applications because it improves discriminating power at around the 0.5 - 1.0 mm level. Also plotted in figure 3.37 are the predicted MTF curves for ideal detectors with pixel sizes of 0.28 mm and 0.56 mm, corresponding to detector pixel sizes of 0.39 mm (1×1 pixel binning) and 0.78 mm (2×2 pixel binning) projected back to the isocentre plane from a source-detector distance of 140 cm. Following the theory in chapter 2, these are sinc functions representing the Fourier Transform of the aperture function described by the square pixel. It is evident from the plots that in neither mode is the detector operating at its full resolving potential, when considering detector element size alone. Blurring caused by scatter in the copper build-up and phosphor layers of the detector, and optical scatter in the phosphor, is clearly significant.

Contrast-to-noise ratio (CNR) results were also calculated from the QEPI1 phantom for all imaging modes, both with and without the licence present. On average, the ratio between full resolution (licence present – 1×1 pixel binning) and half resolution (licence disabled – 2×2 pixel binning) was 0.69 ± 0.02 (1 SD). This was a surprising result – although, as expected, the CNR fell when moving to the finer pixel size it did not fall to 0.5, so is inconsistent with the theory above. Rather, the measured ratio is indicative of the reduction one would expect for a reduction in pixel size by half, leading to a reduction in CNR by a factor of $1/\sqrt{2}$=0.7 . This may be because scatter and optical blurring within the detector are so severe that delivered dose always contributes to more than one pixel, which is consistent with the MTF curves being considerably poorer than those of the ideal detector. Alternatively, it may be the case that the imager electronics do not independently read out every single detector element, so that the noise between pixels is correlated. Also somewhat surprising, but equally important, are the NNPS results in figure 3.38. Whereas these indicate only a slight increase in stochastic noise when the full-resolution licence is enabled, there is also a very significant reduction fixed pattern noise. No definitive explanation has been reached to explain these results. They are interesting findings which warrant further investigation.

From these experiments it appears that whilst operating the Varian aS1000 detector in its 'full resolution' mode there is a significant boost in spatial resol-

ution, there is also an associated reduction in CNR, although not as much as expected by the underlying theory. It therefore may not be the case that the default aS1000 modes of operation may be appropriate for all clinical imaging scenarios and a balance must be struck between resolution and detectability. This is an important result because it is not disseminated to customers by the manufacturer and has not been reported in the literature. Furthermore, basic literature on portal imaging tends to emphasise the importance of detectability and CNR over spatial resolution (see, for example [123, 343]), so this technology is not completely in line with general thinking in the community. This is a very good example of where objective performance metrics can contribute to the optimisation of radiotherapy imaging modalities for particular clinical applications. In Oxford, the possibility of making the reduced-resolution but higher CNR mode available as a routine addition to the standard clinical imaging modes is being considered.

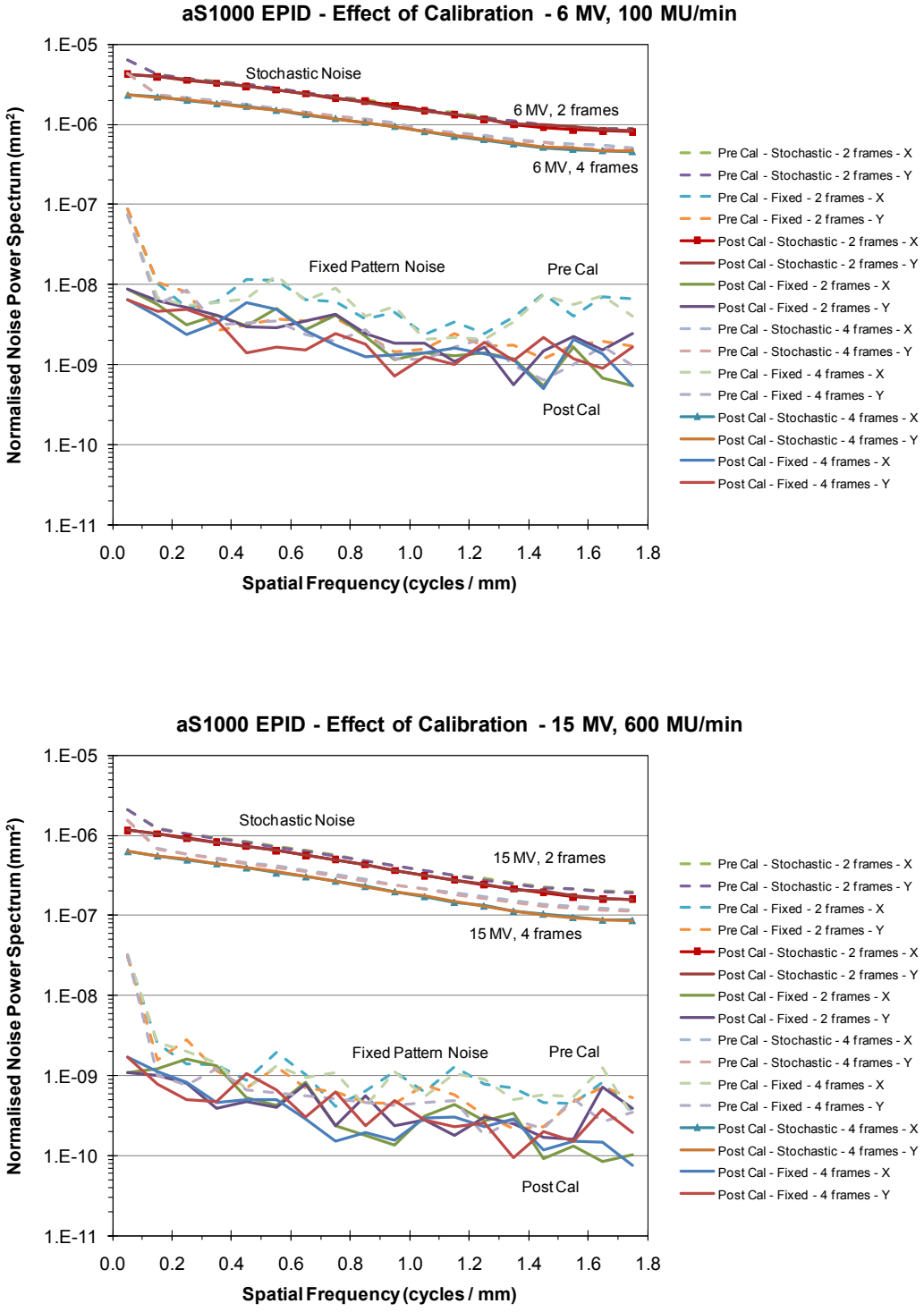


Figure 3.36. – Effect of comprehensive dark and flood-field calibration on Varian aS1000 normalised noise power spectrum (NNPS) for two clinical imaging modes. Top–6MV low dose-rate RadShot. Bottom–15MV high dose-rate RadShot.

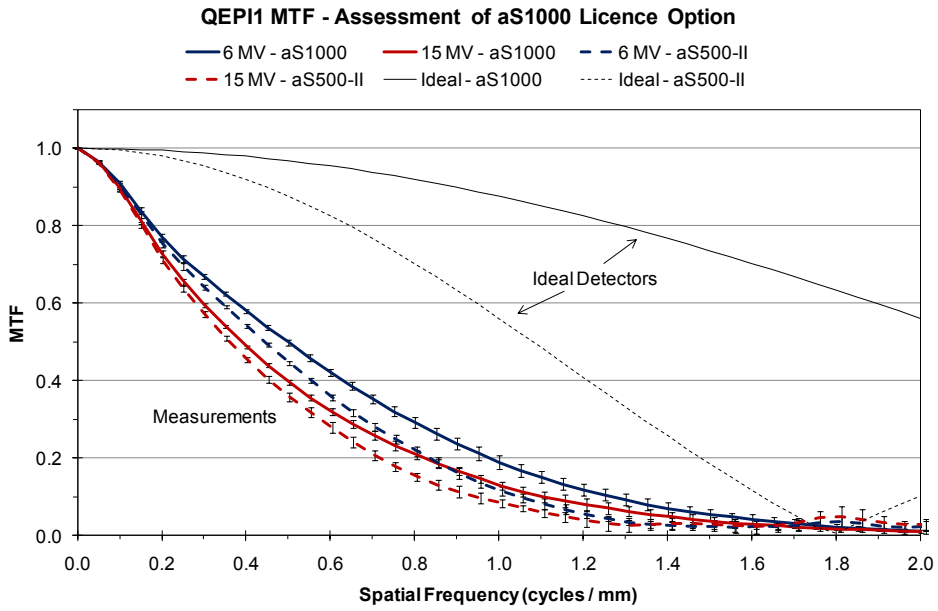


Figure 3.37. – Evaluation of MTF improvement between Varian aS500-II and aS1000 licence options.

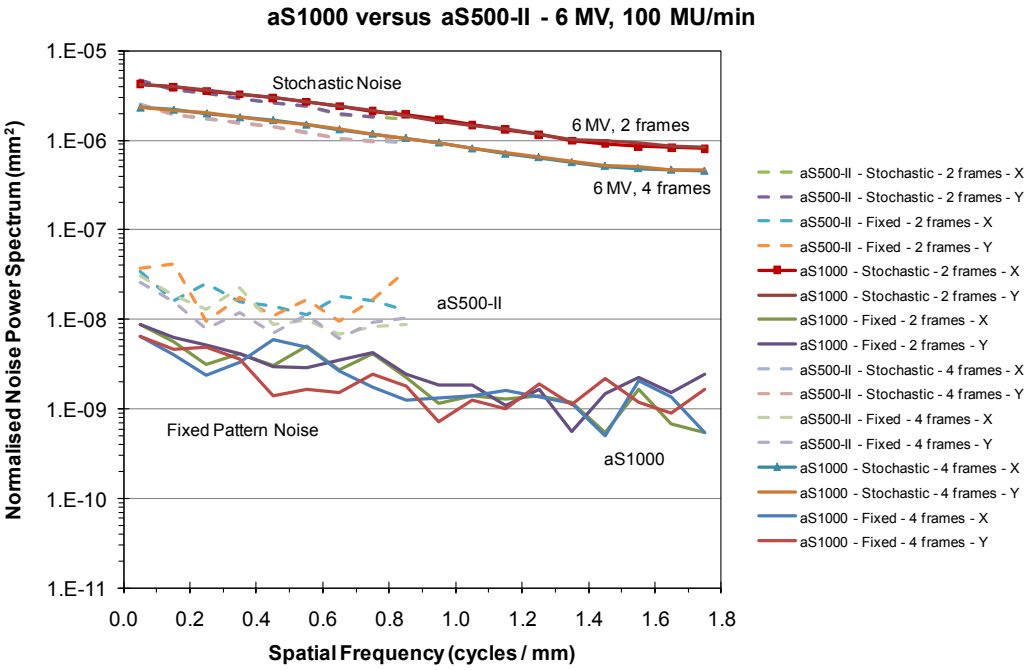


Figure 3.38. – Comparison of normalised noise power spectra (NNPS) of Varian aS1000 and aS500-II detectors. Images were acquired at 6 MV, 100 MU/min, taking an average of 4 frames, at a dose of 2 MU / image.

3.8. Inter-comparison of EPIDs

Throughout the course of this work images were acquired of the QEPI1 phantom using the imaging modes routinely used on all linear accelerators in both Edinburgh Cancer Centre (ECC) and Oxford Cancer Centre (OCC). A detailed list of all linacs was given previously in table 3.1. As described in section 3.2, the routine modes in use are:

- For dedicated imaging fields:
  - Lowest energy, lowest dose-rate, standard quality (2 or 4 frame averages, depending upon detector technology);
- For imaging as part of treatment fields:
  - Low energy, high dose-rate, high quality (4 or 10 frame averages);
- High energy (if available), high dose-rate, high quality (4 or 10 frame averages).

f_{50} values calculated for each linac and mode are presented in 3.39 and the CNR results in figure 3.40.

Reassuringly for relatively new detectors, all OCC f_{50} results for a given energy are the same within the limits of experimental uncertainty, which is taken as ±1 standard deviation of the results measured on the eight QEPI1 edges across the field of view. ECC f_{50} results also agree for any particular energy, with the 6 MV results being lower than those in OCC because the ECC EPIDs were the older aS500 models whilst the OCC EPIDs were the resolution aS1000 devices. This is consistent with the results found when characterising the new aS1000 detector in the previous section. However, an interesting observation is that the f_{50} at 6 MV for 'ECC - LA3' is slightly lower than other linacs with equivalent beams and imagers (ECC - LA1 and LA2) and the error bar on this result is considerably larger. This is due to this EPID suffering from large clusters of defective pixels which were also having a noticeable deleterious effect on clinical image quality. The blurring of the QEPI1 edge which lay on the cluster of pixels therefore resulted in greater variability in measured MTF performance across the field of view. It is very encouraging that IQWorks and the QEPI1 phantom are together sensitive enough to track a clinically relevant localised degradation in spatial resolution. Because of the location of the defects they would not have been detected by the PTW EPID QC phantom, and analysis by the QC-3V phantom was difficult because only a single bar

pattern was affected by the clusters at any time. Nevertheless, simply taking the average QEPI1 f_{50} across the field of view may in itself have been insufficient to highlight this problem. To ensure such problems are detected in the future, the routine analysis protocol for the QEPI1 now includes comparing each of the f_{50} values from the eight edges of the phantom against established baselines.

If the MTF curves are renormalised to unity at 0.1 lp/mm, following the approach generally taken with the QC-3V phantom and PIP Pro, and the f_{50} results recalculated, then the values for the 6 MV aS500 measurements agree with those of other researchers to within 7%[232, 239], which is within the limits of experimental uncertainty.

It is interesting that the 8 MV results for ECC LA4 and LA5 are better than the 6 MV results for the other ECC machines, and that the 15 MV results for these linacs are comparable with those of the linacs in Oxford, when operating at the same energy. This is unexpected because the Oxford linacs have the higher resolution aS1000 detector, which was shown above to offer a tangible improvement in MTF. It is thought this may be due to LA4 and LA5 potentially having a narrower electron beam focal spot, and is another potential reason why the MTF graphs in figure 3.37 were so much poorer than the ideal curves. However, further investigation is required before definitive conclusions can be drawn and unfortunately it is not currently possible to repeat the experiment on these linacs.

Results from the CNR intercomparison were less conclusive. Although certain trends are apparent – such as the 6 MV low dose-rate mode of all OCC linacs having relatively similar performance – there is generally little consistency between linacs and modes. Furthermore, observed differences in performance are well outside the estimated experimental uncertainty, suggesting that other factors are influencing this image quality metric. A deeper investigation revealed that both the range and magnitude of pixel values was considerably different from one EPID in OCC to the next. It is suspected that differences in the internal configuration of the EPIDs (such as underlying gain settings) influence these values and that whilst the devices are calibrated to give good quality images they are not necessarily configured to be consistent with each other. This is essentially a linearisation issue and steps are being taken to rectify it. However, it is an important consideration if consistency in clinical imaging performance from one linac to another is to be assured. Furthermore, it is a known issue which is mitigated for dosimetry applications by following a more involved calibration procedure for designated dosimetric acquisition

modes[347].

A key CNR result worth drawing attention to is the considerably higher value for 15 MV, 600 MU/min images on OCC - V5. This is caused by the contrast measured at 15 MV being almost double that of all other linac / detector combinations operating under the same conditions. Although this may be a linearisation issue, as mentioned above, it is not the whole explanation because this linac behaves differently to the others when acquiring images. Whereas on other linacs the operator hears the periodic 'beeping' of the monitor unit counter slow down as the effective dose-rate is reduced whilst the beam is held off by the imager during frame readouts, on this linac the 'beep' pattern is sporadic bursts. This behaviour is preserved between detector calibration cycles and has not been rectified by attempts to retune the detector's control over pulse modulation. No satisfactory explanation has been reached and this is currently being investigated by the manufacturer.

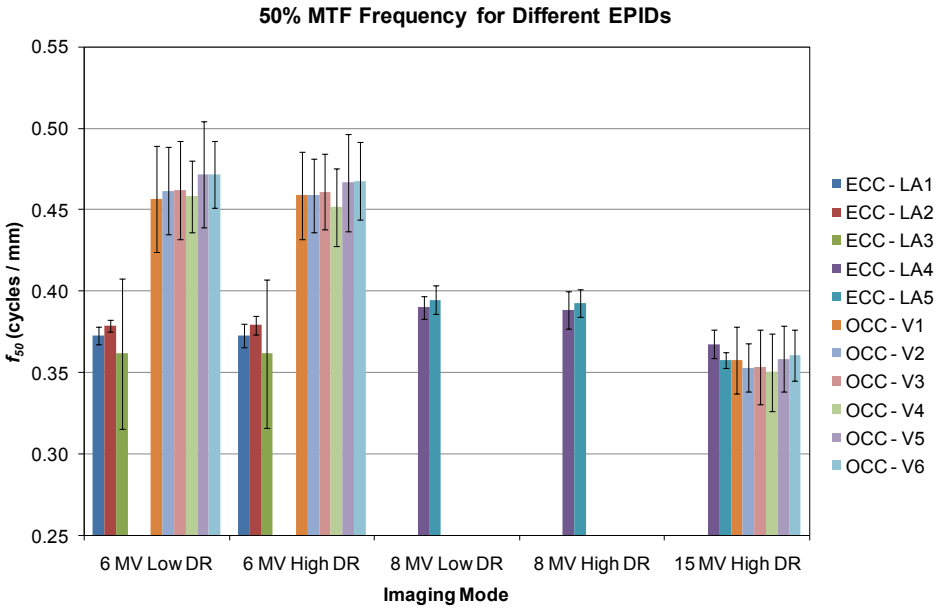


Figure 3.39. – Comparison of f_{50} for different imaging modes and EPIDs. DR denotes 'dose-rate', ECC is Edinburgh Cancer Centre, OCC is Oxford Cancer Centre and the imaging modes and linacs are as described in section 3.2. Error bars are one ±1 SD in measurements made across the field of view.

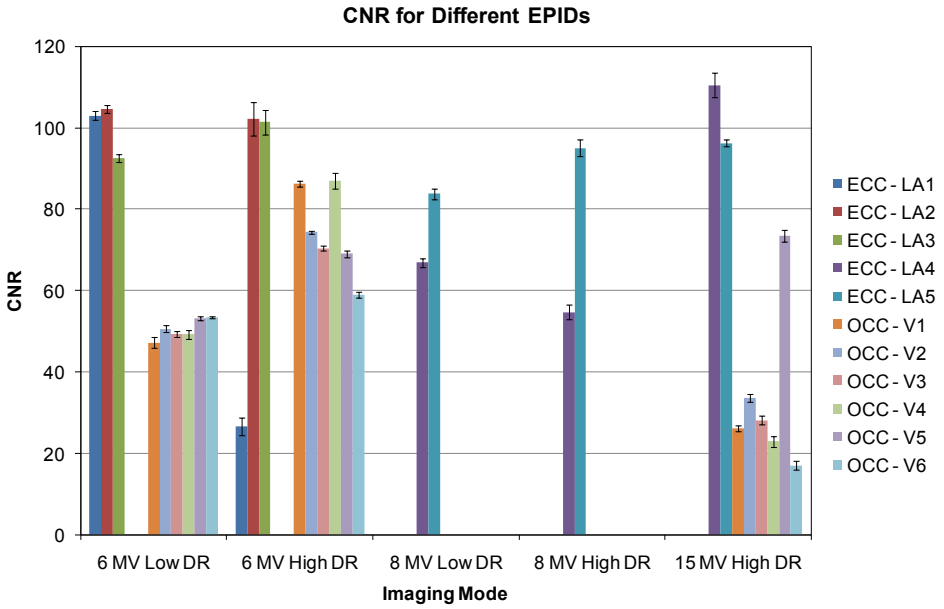


Figure 3.40. – Comparison of contrast-to-noise ratio (CNR) for different imaging modes and EPIDs. DR denotes 'dose-rate', ECC is Edinburgh Cancer Centre, OCC is Oxford Cancer Centre and the imaging modes and linacs are as described in section 3.2. Error bars are ±1 SD in measurements made across the field of view.

3.9. Oxford Cancer Centre EPI QA Programme

Daily imaging of the QEPI1 phantom with immediate analysis by IQWorks forms the core of the EPI quality assurance programme in Oxford Cancer Centre. Although the users performing the daily checks focus on the PASS / FAIL results reported by IQWorks the underlying numerical results are stored to a database to facilitate a more detailed subsequent analysis or the identification of time-trends. Harnessing the Reporting Services built-in to Microsoft SQL Server Express reports on EPI performance are available to physicists at their desktops via a web browser interface. This greatly streamlines the QA reporting process and encourages regular review of results. As an example, the browser interface to the MTF f_{50} results for all 6 MV modes on a single linac are shown in figure 3.41, in which the plot of historical data clearly shows a divergence of results across the FOV when the 5.0 mm prototype QEPI1 jig was replaced with a new 3.5 mm one. (Although the divergence appears very large this is exaggerated by the small scale on the Y-axis. This effect was discussed in detail in section 3.5 above.) Two other examples of the interpretation of historical data are included below.

Also increased transmission through bar patterns at 15 MV[281].

Figure 3.42 shows the time-trend of CNR results for the most commonly used imaging mode on linac 'Varian 4' in Oxford. A clear jump in performance is visible towards the end of 2009. This was due to a spontaneous failure of the EPID and the driver software being reinstalled by the engineer. It is still not fully understood why performance shifted so dramatically following this action and the situation is still being investigated. This may be due to the potential detector configuration differences described in section 3.8 above. However, it emphasises the importance of being able to track performance over time.

Lastly, the time-trend in measured phantom width and height for the same imaging mode is plotted in figure 3.43. It is encouraging that the complete historical data record falls within a very tight band (±1.25 mm) about the mean, which encompasses all uncertainties due to phantom misalignment, errors in detector positioning, etc. Furthermore, the band becomes even tighter ($<$±0.75 mm) with the introduction of a new daily check jig with superior alignment targets.

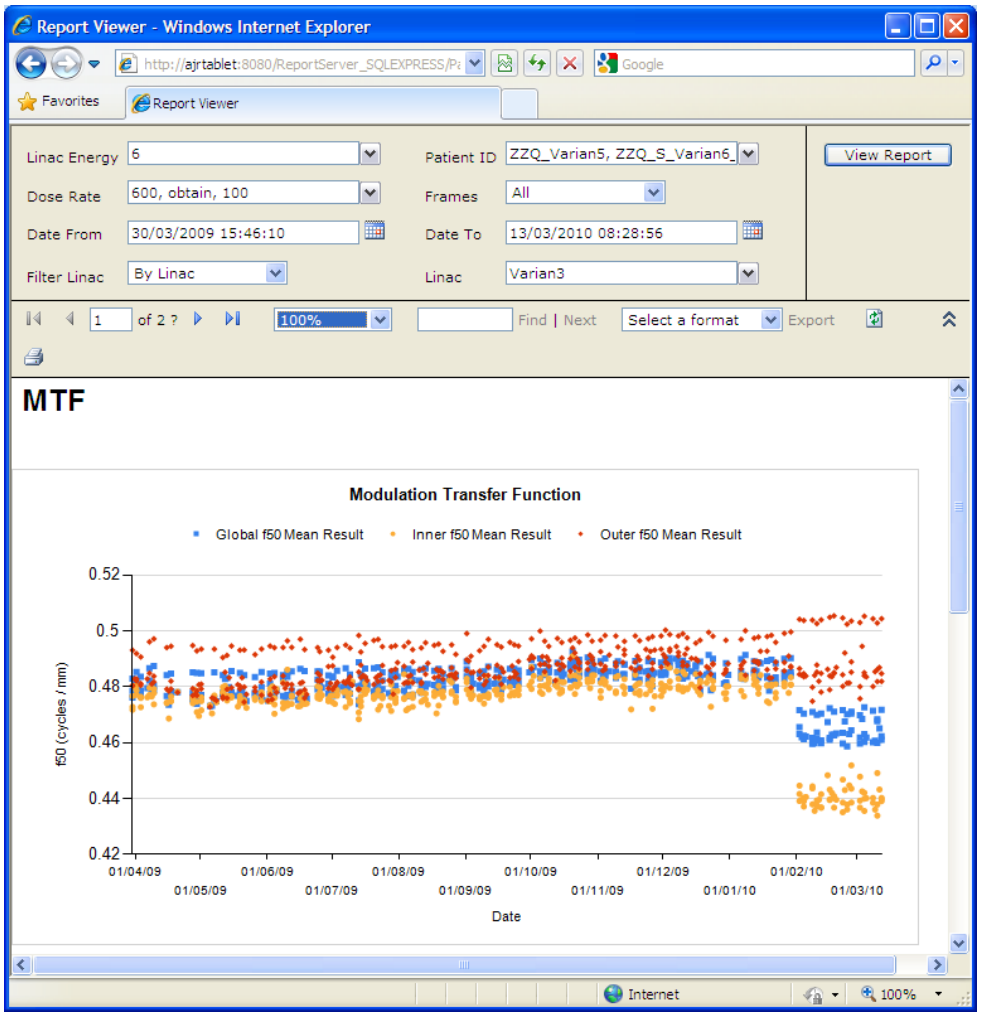


Figure 3.41. – Graphical report of f_{50} time trend data for a single Varian aS1000 EPID, presented in a web browser. Measurements were made daily using the QEPI1 phantom and IQWorks for the 6 MV imaging modes on a single Varian linac.

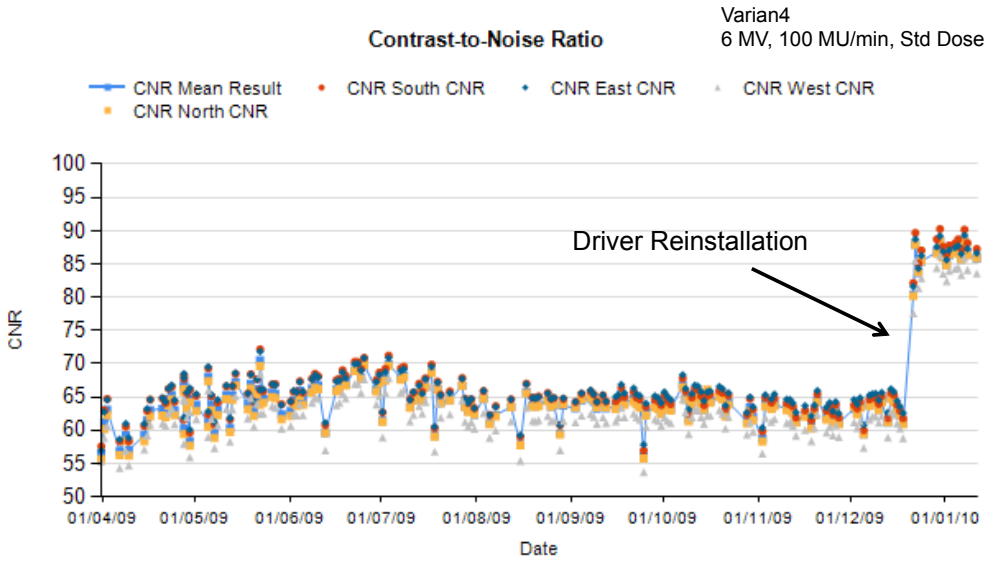


Figure 3.42. – Contrast-to-noise ratio (CNR) time-trend data for a single imaging mode of a Varian aS1000 EPID. A step change in performance resulting from a reinstallation of EPID driver software is clearly visible.

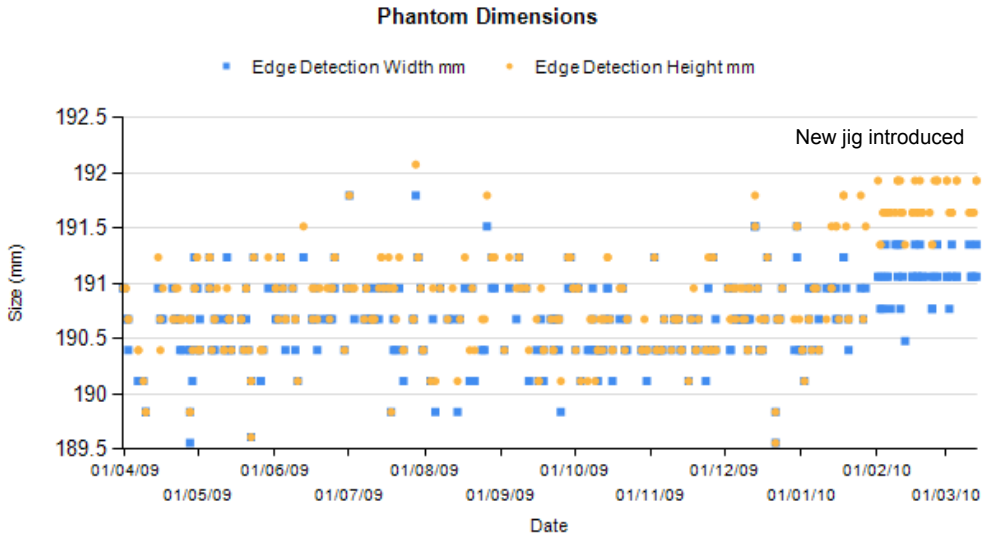


Figure 3.43. – Geometric linearity time-trend data for a single Varian aS1000 EPID, presented as the QEPI1 phantom dimensions calculated daily from 6 MV images by IQWorks.

3.10. Conclusion

Reliable and consistent tools for the objective characterisation of EPID performance are crucial components of any imaging quality assurance (QA) programme or optimisation strategy. It has clearly been demonstrated that IQWorks can be used to analyse images of the QC-3V and PTW EPID QC phantoms using methodologies compatible with the software supplied by the respective phantom manufacturers.

Modulation transfer function (MTF) results between IQWorks and PIPSPro for the QC-3V phantom agree within 1% when the phantom is at isocentre and 2% when it is on the surface of the detector. Furthermore, although PIPS Pro requires a pair of images to be supplied to the MTF calculation algorithm, so that noise correction can be performed using the difference of the two images[281, 381], it was determined that this correction is only essential for the MTF point associated with the highest spatial frequency bar pattern in the phantom. However, it is argued that this is an unnecessary refinement because the PIPS pro implementation of Droege and Morin's methodology[80] is incomplete in that it does not incorporate Coltman's aliasing correction[54]. Instead, for the purposes of performing consistency measurements as part of a routine QA programme it is felt that an analysis based on single images is adequate. This is supported by the results presented in this chapter.

Calculations of other image quality metrics from QC-3V phantom images, including signal, contrast and contrast-to-noise ratio (CNR) are also in excellent agreement between IQWorks and PIPS Pro, to within the limits of experimental uncertainty, although whilst undertaking these measurements it was discovered that the PIPS Pro CNR calculation is not as specified in the literature[281]. Although the approach employed by PIPS Pro is still valid this is an important result because it means that values for CNR reported by other researchers working with this phantom may not be directly comparable, depending upon whether PIPS Pro or different software was used to process the images. An additional benefit of IQWorks was that the analysis scheme could be customised to accommodate defects in the phantom, something not possible in the fully automated commercial software.

Overall, results were not as consistent between IQWorks and the PTW epidSoft software when analysing images of the PTW EPID QC phantom. For both the MTF and signal-to-noise ratio (SNR) assessments this was attributed to unreliable and clearly erroneous decoding of image files by epidSoft. Additionally, the calculation of all but the lowest frequency MTF points by either package

was sensitive to phantom and region of interest (ROI) alignment, with small errors (±1 pixel) in ROI placement resulting in changes in calculated MTF of up to 40% in IQWorks and 150% in epidSoft. In part, the intrinsic uncertainty in these measurements is due to design flaws: the highest spatial frequency patterns are orientated to align with the pixel matrix, so that slight translations or rotations in the phantom will result in interference artefacts which will have a significant impact on measured variance, and these patterns which are the most sensitive to alignment issues also occupy the smallest surface area of any of the bar patterns, thereby making them the most difficult to align ROIs with correctly. However, it was also demonstrated that the ROIs employed by epidSoft are over large, intended to cover the full extent of a bar pattern under perfect alignment conditions, so that even small misalignments will result in ROIs extending beyond the edges of their patterns and resulting in overestimations in the measurements of local variance.

Better agreement was observed between IQWorks and epidSoft for measurements utilising ROI-based mean signal measurements. Calculations of contrast, signal linearity – across the extended field of view – and local signal linearity were in agreement between the packages to within the limits of experimental uncertainty. However, it is noted that the epidSoft approach of rescaling linearity results to lie between the measured values on the first and last steps in the attenuation patterns causes a great deal of useful information to be lost: differences between beam energies are suppressed and local variations in system transfer properties across the field of view cannot be identified. Interestingly, the transmission of the step wedge which should nominally be 50% at 6 MV was found to be only 38% at this energy.

Given the number of design and implementation issues encountered it is difficult to be confident that epidSoft and the PTW EPID QC phantom could together reliably form the foundation of an optimisation strategy or QA programme. Certainly, epidSoft's file handling issues make it extremely difficult to undertake comparative performance evaluations between equipment from different manufacturers, where the ranges of pixel values present in the images might be considerably different. However, some of these issues are resolved by IQWorks' demonstrably more robust file handling. Also, the flexibility afforded by being able to adjust the location and size of individual regions of interest in IQWorks enables other limitations of the phantom to be addressed.

In this chapter the new QEPI1 portal imaging phantom was introduced. Being straightforward and inexpensive to manufacture the QEPI1 phantom is readily

accessible to all radiotherapy departments. Six instances of the phantom were constructed using a simple manufacturing process and the calculated performance metrics were found to be consistent across all jigs. This demonstrates the ease with which this new phantom might be cheaply and efficiently adopted by different departments, with it being possible to modify the exact dimensions and mounting arrangement to meet local requirements.

When comparing the different acquisition modes of the aS1000 detector it was found that there was a slight boost in contrast between low and high dose-rates. This was surprising because the beam spectrum and irradiation conditions across the field of view should be the same regardless of dose-rate. Although small, the boost was observed over repeated measurements and appears to be outside the range of experimental uncertainty. Comparisons of the CNR for the aS1000 'high quality' (4 frame averages) and 'standard dose' (2 frame averages) modes indicated that CNR is proportional to $\sqrt{dose}$, in line with the underlying Poisson statistics of the system. Knowing that full benefit may be gained by increasing the number of frames utilised in an exposure allows decisions to be made regarding what level of dose may be required for a specified improvement in CNR, which may be useful for some clinical applications and the development of new imaging protocols.

MTF curves calculated with the QEPI1 phantom agreed well with those derived from images of the QC-3V and PTW EPID QC phantoms when these images were processed by IQWorks and following the fundamental theory on which bar pattern MTF analysis is based[54, 80], rather than the implementations of this theory in the commercial software bundled with the phantoms[275, 381]. These results are encouraging because they essentially validate the new phantom and support the use of edge-based MTF methods for EPID performance evaluation by the wider radiotherapy physics community. In addition, being able to demonstrate that IQWorks can yield compatible results from three independent phantoms indicates that objective performance comparisons between different centres can reliably be performed using whatever phantoms are locally available, as long as care is taken over the analysis algorithms employed. The IQWorks framework developed in this thesis is the ideal tool for this. Furthermore, if the goal is to evaluate new detector technologies in the context of other devices previously characterised in the literature, IQWorks can also be utilised to reconsider images acquired in the past, but this time applying the new analysis methods. Similarly, this approach opens up the possibility of comparing the performance of radiotherapy imaging equipment

with results widely reported in the diagnostic imaging literature, and vice versa.

Examples of the QEPI1 being applied to detailed characterisation experiments, to address optimisation questions and as part of a routine QA programme have been discussed throughout this chapter. Key advantages of the QEPI1 over other phantoms are identified, including its ability to calculated metrics both locally and across the whole field of view, its simple design making it straightforward and inexpensive manufacture, and its methods being underpinned by current best practice in the diagnostic imaging field. This phantom presents an opportunity to establish performance baselines across multiple centres which can be utilised in widespread clinical optimisation exercises.

An interesting result was how the MTF was found to vary across the field of view, being marginally poorer closer to the beam central axis. Further investigation is required, possibly involving Monte Carlo modelling, to explain the physics of this fully, as well as to determine any implications for clinical imaging.

Other important findings were the differences in performance measured between Varian's top of the range aS1000 detector, and its less expensive aS500-II device. It was determined that, as claimed and expected due to the underlying physics, the smaller pixel size of the aS1000 unit results in a broad spectrum improvement in MTF. In particular, there was improvement of between 50% and 100% over the clinically relevant range of 0.4–1.4 cycles / mm. This could prove important for applications requiring sub-millimetre image guidance, such as stereotactic radiosurgery. However, and also as expected, the improved spatial resolution comes at the expense of CNR, although the reduction by $\sqrt{2}$ was not as pronounced as the factor of 2 predicted. Nevertheless, this result is significant because it is not widely publicised yet clinical issues with megavoltage portal imaging most commonly arise from images being inherently noisy and of low contrast. In its default configuration the aS1000 licensed detector cannot be run in the 2×2 binning mode even though this may be advantageous for certain clinical applications. It is suggested that Varian might consider providing this facility as an option, giving clinical users the choice between a high resolution, low CNR mode and a lower resolution, high CNR mode at the same dose point, thus enabling the detector to be utilised to its full potential.

Through performing an intercomparison of clinical imaging modes across eleven linacs in two radiotherapy centres, the flexibility and potential of the new QEPI1 phantom and the implementation of the analysis framework in IQWorks were clearly demonstrated. Results were generally consistent with

expectation, with clustered pixel defects on one detector being highlighted through wider than expected variations in MTF across the field of view, and anomalous CNR measurements for one imaging mode of another detector meriting further investigation.

Traditionally, quality control measurements undertaken as part of a routine QA programme tend to be fairly simplistic PASS / FAIL tests utilising qualitative or semi-quantitative phantoms. This is especially true for those tests performed frequently and where there are economic and time pressures to release equipment back into clinical use. However, by integrating the QEPI1 phantom into the daily check process across all six linacs in Oxford Cancer Centre the feasibility of performing fully objective and quantitative measurements on a regular basis is clearly demonstrated. A full analysis involving MTF calculation across the field of view, geometric linearity, system transfer properties and CNR is performed every morning on two clinical imaging modes on every linac, with the numerical values being automatically compared against baselines then flagged as simple PASS / FAIL results. This has been widely accepted by the technical staff involved in undertaking the measurements and is recognised as incurring only a minimal time penalty. However, the distinct advantage of this approach over traditional methods – including the temptation to analyse resolution patterns in test objects like the QC-3V and PTW EPID phantom by eye, or counting contrast-detail objects in the PTW EPID phantom or standard Varian phantoms – is that because the numerical values are unambiguous and operator independent they can be utilised to accurately chart and identify performance trends. Furthermore, in the event of an equipment problem an experienced operator can look beyond the reported PASS / FAIL summary results at quantities which more deeply describe the fundamental operation of the system. Finally, a clear advantage of this approach discovered early in its implementation in Oxford is that even following very significant maintenance on a portal imaging system the metrics provided through the standard daily test are sufficient to verify the equipment is suitable for clinical use.

It is strongly suggested that objective image quality metrics already well established in diagnostic imaging are increasingly introduced into the radiotherapy sphere, complementing – and where appropriate replacing – some of the more traditional approaches.

Chapter 4.

Modality 2: X-Ray Computed Tomography – Fan-Beam and Cone-Beam

4.1. Overview

X-ray Computed Tomography (CT) is widely used as the reference modality against which radiotherapy treatment plans are prepared. As described in chapter 1, this modality's good soft-tissue contrast, high degree of spatial uniformity, pixel values with a stable, reproducible relationship to fundamental electron density, millimetre or sub-millimetre level spatial resolution and extremely robust geometrical accuracy are all import factors for tumour localisation, decisions about target and avoidance volumes and the development of an appropriate treatment plan. Increasingly, CT is also now being employed to verify patient set-up at the treatment unit, with the volumetric information provided allowing a more accurate comparison against the simulation CT scan than planar images and enabling a better understanding of the effect of patient set-up corrections on dose delivery(see, for example [161, 337, 358]). Furthermore, there is potential to recalculate or adjust the treatment plan based on the new CT datasets. CT is therefore a key imaging modality throughout the entire radiotherapy patient pathway[155, 250].

Traditional CT scanners employ a narrow fan-like X-ray beam to irradiate a thin 'slice' of the patient. Opposite the X-ray tube, and moving in tandem with it whilst it rotates around the patient, is a narrow row of detector elements which measure the intensity of X-rays transmitted across the fan-beam at each rotation angle. A cross-sectional image of each slice is then reconstructed using simple filtered back-projection algorithms which assume that photons attenuated from

the beam due to scattering interactions and photoelectric absorption do not contribute to the signal at the detector[121, 131, 168]. Advances in detector technology, along with faster signal processing pathways, enable multiple lines of detector elements to be joined together so that more than one slice can be acquired simultaneously[153]. In a single X-ray tube rotation a longer volume can be imaged using thinner slices, and such 'multi-slice CT' (MSCT) allows more sophisticated scanning protocols to be developed, such as extended volume breath-hold sequences[319], dynamic '4D' scans[208] and thin slice whole body scans which are not subject to motion artefacts[190]. X-ray tube technology has also developed in parallel, with more stable tubes facilitating faster rotations and hence quicker scan times[202], and higher capacity tubes enabling both increased patient throughput[164] and the ability to maintain an acceptable signal to noise ratio (SNR) over an extended volume of thin slices. This allows the full potential of the advanced detectors to be realised.

As the number of slices increases the width of the fan-beam must also increase to cover the correspondingly larger detector array. However, as the fan-beam becomes wider and more cone-like the assumption that scattered photons do not contribute to the detector signal begins to break down, and more sophisticated reconstruction algorithms are required, such as that developed by Feldkamp, Davis and Kress (the FDK algorithm)[96].

With the recent widespread availability of large-area flat panel detectors for general radiographic imaging, researchers have begun to utilise these for CT, effectively providing hundreds of rows of detector elements and taking the cone-beam geometry to an extreme. Complex, often iterative or model-based corrections are now required to compensate for the significantly increased contribution of scatter to the signal at the detector[362]. Although commercial implementations are available this is still an evolving technology, with image quality from flat panel based system generally poorer than that of traditional scanners. Because of this, and also as a result of the special requirements of the reconstruction algorithms, a distinction tends to be made between the two technologies. The acronym 'CT' implicitly refers to X-ray CT using either a fan-beam, or in the case of multi-slice scanners, a minimal cone-beam geometry, whilst 'CBCT' designates the use of an extended cone-beam geometry with a flat panel detector.

Both CT and CBCT are widely used for radiotherapy applications. CT is still the modality of choice for pre-treatment imaging due to its superior image quality and lower patient doses, although conventional radiographic simulators

with a CBCT option are available because their greater clearance around the patient offers advantages for some applications[223]. However, with wide-bore CT scanners becoming more widespread the future of the CBCT simulator is uncertain.

CT and CBCT options are also available for verifying patient set-up immediately prior to or following treatment. In this scenario CBCT is more practical because an X-ray tube and flat panel detector can easily be attached to the linac gantry at 90° to the megavoltage beamline, leaving the treatment room layout and services unchanged[161, 337, 358]. However, CT-on-rails solutions are also available which involve a conventional CT scanner gantry moving along the couch to acquire a scan of the patient in the treatment position. Although more cumbersome, these offer all the advantages of a standard CT scanner[56, 192]. Megavoltage X-ray CT (denoted 'MVCT') is also possible[273] and as an extreme case the Tomotherapy system[211] performs CT scanning for treatment planning and set-up verification using a similar megavoltage fan-beam as for treatment delivery.

This chapter explores the application of the IQWorks image analysis framework to CT with emphasis on the image quality factors important for radiotherapy. Two radiotherapy specific CT phantoms are introduced – the RMI 467 (Gammex, Wisconsin, USA) and the Varian Performance Phantom (Varian Medical Systems, Palo Alto, California, USA) – and analysis trees are developed to analyse the images from each of these, calculating electron density calibration curves, CNR, modulation transfer function (MTF) and geometric linearity. Investigations are described examining the effect on these metrics of varying X-ray tube accelerating potential (kVp), tube current (mA) and whether helical or axial acquisition modes are employed.

Although only recently applied to radiotherapy applications[293, 298], the Catphan series of phantoms (Phantom Lab, Salem, New York, USA) has been in widespread use in the diagnostic imaging community for many years. Detailed IQWorks analysis trees are constructed for each Catphan module, calculating electron density calibration curves, CNR, MTF, geometric linearity and metrics of uniformity. Similar experiments were performed as with the traditional radiotherapy phantoms and comparisons drawn between the results. In particular, the suitability of the Catphan as a general purpose radiotherapy phantom is evaluated.

Between them, these phantoms form the basis of the CT quality assurance programmes in Edinburgh and Oxford. Many months worth of sensitometry,

CNR, uniformity and geometry results are presented and significant shifts or spikes in performance trends identified and discussed.

During this study two linacs equipped with Varian's on-board imaging (OBI) CBCT facility[344] were installed in the new Oxford Cancer Centre, along with a new Acuity conventional simulator with a CBCT option[346]. The analysis techniques introduced for conventional CT scanners were applied in the detailed commissioning and initial optimisation of these units, facilitating performance comparisons against conventional scanners. All clinical scanning modes are considered and the influence on image quality metrics of varying acquisition parameters such as field of view and matrix size are investigated in detail. In addition, the normalised noise power spectrum (NNPS) is calculated for the OBI system and compared against that of a conventional CT scanner for an equivalent patient scanning protocol.

4.2. CT Systems Considered in this Study

Five CT units were considered in this work: a conventional single-slice scanner in Edinburgh, a conventional multi-slice scanner in Oxford and three CBCT systems, all in Oxford. Technical specifications of each of these units are included in table 4.1. Of the two conventional scanners the model in Oxford (HiSpeed QX/i) is essentially a next generation version of the unit in Edinburgh (HiSpeed FX/i). All three of the CBCT systems utilise the same acquisition and reconstruction technologies. However, two of these are Varian's On-Board Imaging (OBI) system, which is intended for verification of patient set-up during treatment and the other is a Varian Acuity simulator, which acquires CBCT images for treatment planning. The two OBI units are associated with linacs V1 and V2 in table 3.1 in chapter 3.

Whereas the dimensions of the scan aperture, the X-ray source-isocentre distance (SID) and source-detector distances (SDD) are all fixed on the conventional CT scanners, there is flexibility in the CBCT systems for the scan geometry to be adjusted. Varian's OBI solution consists of a kilovoltage X-ray tube and a flat panel detector mounted opposite each other at 90° to the treatment beamline, both mounted on robotic arms which allow them to be moved out of the way during patient set-up or treatment and which also provide considerable flexibility in SID and SDD. In the Acuity, only the detector is mounted on a robotic arm, whereas the X-ray tube is fixed at 100 cm SID. Being able to modify SID and / or SDD facilitates optimisation of the acquisition geometry for particular treatment

Model	HiSpeed FX/i	HiSpeed QX/i	OBI 1.4	Acuity 8.5
General Details				
Centre	Edinburgh	Oxford	Oxford	Oxford
Manufacturer	GE	GE	Varian	Varian
Technology	CT	MSCT	CBCT	CBCT
Acq Modes	Axial, Helical	Axial, Helical	Axial	Axial × 3
SID (cm)	54	54	100*	100*
SDD (cm)	95	95	150*	150, 160*
Aperture (cm)	70	70	Variable	Variable
References	[45, 149]	[102, 152, 175]	[344]	[346]
X-ray Tube				
Heat Capacity (MHU)	3.5	6.3	2.0	2.0
Potentials (kVp)	80, 120, 140	80, 100, 120, 140	100, 110, 125*	125*
$\tilde{E}$ **(keV)**	48, 60, 65	48, 54, 60, 65	50, 53, 57	63
Focal Spot (mm × mm)	Small: 0.6 × 0.6 Large: 1.2 × 1.2	Small: 0.7 × 0.6 Large: 0.9 × 0.9	Small: 0.4 × 0.6 Large: 0.8 × 1.1	Small: 0.4 × 0.6 Large: 0.8 × 1.1
Filtration (mm Al)	5.9	5.7	2.7	7.6
Rotation (s)	1, 1.5, 2, 3	0.7, 1, 2, 3, 4	40, 70	70
Detector / Acquisition				
Name	HiLight	HiLight Matrix	PaxScan 4030CB	PaxScan 4030CB
Technology	Ceramic	Ceramic	CsI(Tb) / aSi	CsI(Tb) / aSi
Elements: tran	793 + 23	880 + 32	1024	1024
Elements: lng	1	16	512	512
Det size (mm) : tran	580	518	400	400
Det size (mm): lng	1.25	20	300	300
Element size: trans (mm)	0.71	0.57	0.39	0.39
Element size: lng (mm)	1.25	1.25	0.39	0.39
Views / s	972	1408	15	15
Projections	Variable	Variable	650	650
Reconstruction				
Max SFOV (cm)	50	50	25.5, 45	25.5, 45
Max DFOV (cm)	50	65	25.5, 45	25.5, 45
Recon Matrix	512^2	512^2	128^2, 256^2, 384^2, 512^2	128^2, 256^2, 384^2, 512^2
Max slices	1	4	165	165
Slice thickness (mm)	1, 2, 3, 5, 7, 10	0.625, 1.25, 2.5, 3.75, 5, 7.5, 10	1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 10	1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 10
Recon Filters	Standard, Standard+, Soft, Detail, Lung, Bones, Edge, Performance	Standard, Soft, Detail, Lung, Bones, Bones+, Performance	Standard, Sharp	Standard, Sharp
Post Processing	Numerous, noise and artefact reduction	Numerous, noise and artefact reduction	Limited, ring artefact suppression	Limited, ring artefact suppression

Table 4.1. – Specifications of CT systems considered in this study. *denotes default clinical values which can be modified. SID = Source-isocentre-distance. SDD = Source-detector-distance. lng = longitudinal direction (i.e. patient long axis). trans = transverse direction (i.e. parallel to scan plane). $\tilde{E}$ = effective energy. MSCT and CBCT refer to multi-slice CT and cone-beam CT respectively. SFOV = scan field of view. DFOV = display field of view. Det = detector. Detector element dimensions are as projected back to the isocentre plane.

sites. For example, certain breast treatment setups require a very large aperture to accommodate a patient on an angled board and with an elbow extending outwards from the couch, whilst only a small aperture may be necessary for a brain scan. According to Poisson statistics, image noise is inversely proportional to the square root of dose to the detector so that keeping the noise below a defined level requires a minimum dose to the panel. All other considerations being equal, the shortest SDD is always preferable so that the X-ray technique settings – and hence the dose burden to the patient – can be minimised. Being able to adjust the SID and / or SDD therefore allows the operator to balance patient dose, image quality and ease of patient accommodation for particular treatment sites. However, each acquisition geometry requires its own dedicated calibration, a relatively onerous process involving the scanning of multiple phantoms, and would need to be regularly tested as part of the ongoing quality assurance (QA) programme. Therefore, despite the potential advantages of variable SID and SDD, the tendency is to configure a few standard modes for use across all treatment sites. These are discussed later in section 4.3.

Another key difference between the CT systems examined is the range of available fields of view (FOVs). In this context, the 'scan field of view' (SFOV) refers to the diameter of circle over which projection data are collected, whereas the 'display field of view' (DFOV) is the diameter of circle of the reconstructed image. For all systems, the maximum SFOV and maximum DFOV were equal, except for the HiSpeed QX/i which incorporates a 'wide FOV' facility that extrapolates acquired data to extend the standard 50 cm SFOV to 65 cm. This is advantageous for the treatment planning of large patients, or for treatment sites such as breast where the patient set-up and immobilisation devices cause the volume of interest to be near the edge of the scanner bore. However, the data points beyond the SFOV acquisition circle are estimated by modelling transmission through the patient and determining the most likely patient section for the projections measured. These algorithms can never be completely accurate because they essentially calculate a 'best guess' from missing data, and measurements have shown that contrast-to-noise ratio (CNR), geometric linearity and pixel values deteriorate beyond the 50 cm SFOV. On both conventional CT scanners a wide range of FOVs are available, up to the maximum of 45 cm.

Two SFOVs are available on the Acuity and OBI systems: 25.5 cm and 45 cm. These correspond to so-called 'full-fan' and 'half-fan' acquisition modes, referring to the nature of the bow-tie filters that are mounted to the faceplate of the X-ray tube to compensate for varying transmission across the width of

a typical patient section. During a full-fan acquisition the flat panel detector is centred on the X-ray tube central axis, the field of view is limited to 25.5 cm and the gantry rotates through 200°. Alternatively, during a half-fan scan the panel is offset laterally and the X-ray tube collimation adjusted accordingly so that only half the patient is irradiated at any gantry angle. In this geometry the FOV can be up to 45 cm, but a full 360° rotation is required so that projections can be collected along sufficient ray lines through the patient[71, 344, 345].

All the CT units have X-ray tubes with similar focal spots sizes. However, the Acuity and OBI both utilise standard X-ray tubes, whereas the conventional CT systems employ tubes designed specifically for CT scanning. In particular, the HiSpeed QX/i tube has a relatively high heat capacity to facilitate long scans of thin slices in a timely fashion, whilst maintaining a good patient throughput. During clinical use it is rarely necessary to have to wait for either of the conventional CT systems to cool down before continuing to scan, whereas both the OBI systems and the Acuity are limited to around five high quality (i.e. high dose) scans per hour[344].

Each X-ray tube can operate at a range of accelerating potentials (kVps). As the kVp increases the beam spectrum becomes more penetrating, allowing thicker or more dense sections to be imaged. However, at the same photon fluence patient dose generally increases with the square of kVp[200] so increasing kVp whilst maintaining all other scan parameters the same results in a higher dose burden to the patient.

Computed tomography reconstruction algorithms are based on the premise that a two dimensional linear attenuation map can be constructed from transmission measurements at different angles. However, the linear attenuation coefficient of any given material is a function of energy so that an accurate attenuation map can only be deduced either by using a monochromatic beam of photons or by incorporating sophisticated compensatory and sometimes iterative algorithms into the reconstruction process. A particular problem is that photons lose energy as they Compton scatter within the patient so that the energy distribution of the beam changes as it passes through the patient, with the attenuation coefficients also changing accordingly. One way of mitigating this is to utilise additional tube filtration to harden the X-ray beam prior to its impinging upon the patient. By preferentially attenuating low energy photons so that the average energy present in the beam spectrum increases the relative change in energy with distance through the patient is reduced. This is why the two conventional CT scanners and the Acuity all have relatively high total filtra-

tions (of the order of 6-7 mm Al). In contrast, the OBI filtration of 2.7 mm Al is comparable with that of a standard radiographic X-ray set and is made possible because Varian's OBI reconstruction algorithm incorporates model-based beam hardening corrections[345].

It is worth noting here that the linear attenuation coefficients of materials commonly found in the body are linked by a well-defined relationship to their electron densities, information which facilitates dose calculations in radiotherapy treatment planning[18, 145, 242]. Determining the calibration curve between CT pixel value and electron density is discussed in section 4.4. However, it follows that if the measured map of linear attenuation coefficients is a function of kVp, then so is the electron density calibration curve. A caveat is that the significant beam hardening caused by the additional filtration means that the numerical magnitude of kVp is not in itself sufficient to predict the electron density calibration curve of a particular scanner. Instead, the 'effective energy', $\tilde{E}$ may be a more useful indicator and is the average energy of all photons in the beam spectrum. Values of $\tilde{E}$ for the kVp settings of the scanners considered in this study were calculated from kVp and filtration data using the IPEM Report 78[58] Spectrum Processor software[291] and are included in table 4.1.

Three different detectors are used across the CT systems, all proprietary to their manufacturers. GE 'HiLight' detectors comprising ceramic scintillators are utilised on the conventional CT systems, with the multi-slice scanner employing a 'matrix' version which bonds together 16 lines of detector elements, as well as more densely packing the elements along each line. Although there are 16 lines in the matrix the readout electronics are only capable of returning 4 channels of data simultaneously, allowing a maximum of 4 slices to be acquired at a time. However, detector lines may be grouped to return slices of different thickness (e.g. 4×1.25 mm slices, or 4×5 mm). Either side of the imaging portion of each HiLight detector are additional elements used for calibration and normalisation purposes. (This is the second set of figures for the number of elements in the transverse direction listed in table 4.1.) Each detector array is mounted along an arc, so that the distance between each detector element and the X-ray source is constant.

In comparison, the OBI and Acuity units employ Varian PaxScan 4030CB flat panel detectors where the diverging beam geometry means that the path length from the X-ray source to any detector element increases with distance from the beam central axis. These detectors are large-area matrices of amorphous silicon

photodiodes with a CsI(Tb) phosphor. They are optimised for CBCT applications with a high sensitivity, high dynamic range and low veiling glare[21, 225]. The PaxScan 4030CB can operate in a number of modes:

- **Single-pulse, Full-Resolution (SPFR).** Every element of the 2048 $\times$ 1536 detector matrix is read out individually following a single X-ray exposure. This yields the highest resolution images, but also those with the lowest contrast-to-noise ratio (CNR) at a given dose point because the pixel size is smallest. This mode may be used by the Acuity and OBI for high resolution radiographic imaging.

- **Single-pulse, Half-Resolution (SPHR)**. Similar to SPFR, except that the pixels are binned 2 $\times$ 2, reducing the matrix to 1024 $\times$ 768 and with a corresponding reduction in spatial resolution. However, the readout time is less and there is a corresponding boost in CNR (theoretically by a factor of 2) due to the larger pixel size. This mode may be utilised for lower resolution, higher CNR radiographic images or for near 'real-time' fluoroscopic imaging on Acuity and OBI units. It is studied in the following chapter, in section 5.2.

- **Dual-gain, Half-Resolution (DGHR).** Similar to SPFR, except that the pixels are binned 1 $\times$ 2. i.e. in groups of 2 along each row[225]. Following an exposure the panel is read out with high and low gain values being applied to alternate rows. If a detector element is found to have saturated, then the signal from the same pixel in the adjacent lower gain row is chosen; if it has not saturated, then its own signal contributes to the image. This mode effectively extends the dynamic range of the detector by intelligently boosting signal in regions of low exposure and high noise, but suppressing signal in regions of high exposure and saturation. However, it does so at the expense of spatial resolution and CNR. It is used primarily for CBCT acquisitions, where the panel may be exposed to a wide range of doses at any projection angle, but can also be manually selected on the Acuity and OBI for radiographic exposures. (NB: In table 4.1 the detector is specified as having a matrix of 1024 $\times$ 768 because of the 1 $\times$ 2 pixel binning with row interleaving.)

- **Dual-pulse, Half/Full Resolution (DPHR/DPFR).** These are equivalent to SPFR and SPHR, but aim to extend the dynamic range of the detector through a different mechanism from DGHR. Instead of applying different

gain settings to alternate rows after a single exposure, two images are acquired in rapid succession at two exposure settings[224]. If possible, the signals from pixels in the higher dose image are chosen, but if they have saturated then the signals from the lower dose image are taken instead. These modes potentially achieve the same spatial resolution and CNR as SPFR or SPHR, whilst minimising saturation and low exposure artefacts, but with a patient dose penalty. This mode is available only on Acuity units and not on OBIs.

Fundamentally, all the CT systems operate in a similar manner: the X-ray tube and detectors rotate around the patient about a common axis, with the detectors acquiring projection data whilst the X-ray tube exposes. However, whereas the GE HiLight detectors sample at a constant rate with the X-ray tube exposing continuously, the Acuity and OBI projections are acquired at discrete gantry angles. Although the gantry rotates at a constant angular velocity the X-ray tube pulses regularly throughout the rotation so that projections are acquired at a rate of 15 frames per second. Another difference is that whilst slip ring technology permits the X-ray tube / detector assembly of the conventional CT scanners to rotate continuously, acquiring slice after slice as the patient advances through the gantry, the Acuity and OBI gantries perform only a single revolution at a static couch position, imaging a volume of length around 14–17 cm (depending upon SID and SDD). For verification imaging prior to treatment this is usually sufficient because the volume being treated should ordinarily be at the centre of the scan. However, it is not generally sufficient for treatment planning purposes so the Acuity includes a facility to acquire up to three cone-beams consecutively – automatically moving the couch in between each exposure then the gantry rotating back in the opposite direction – then stitching the irradiated volumes together to yield a longer stack of images. Also, whereas the conventional scanners can acquire slices using second or sub-second rotations, allowing motion artefacts to be suppressed or overlapping datasets to be quickly acquired which track motion over time and can be rebinned into 4D sequences, the considerably slower rotation times of the Acuity and OBI result in motion being smeared out.

It is difficult to make informed comments about or to predict the impact on image quality from the sampling information presented in table 4.1. Although, like all modalities involving ionising radiation, relative noise decreases with the square root of dose, spatial resolution is determined by a complex relationship between detector element size, angular sampling rate and the capabilities of the

reconstruction algorithm. Therefore, although the detector element size for the conventional CT scanners is larger than those in the Acuity and OBI it does not automatically follow that spatial resolution in a scan will be poorer. Indeed, the considerably higher sampling rate, higher performance detectors which have been developed explicity and exclusively for CT applications and shorter SDD, may permit a greater number of projections at a higher angular sampling density and with lower noise at the same dose point. Furthermore, image quality metrics – particularly modulation transfer function – are heavily dependent upon the convolution filter, data pre- and post- processing and other details of the reconstruction algorithm[168, 169]. Many more years of development have been invested in the reconstruction algorithms of conventional CT systems than for CBCT, and the reconstruction problem is itself simpler, not having to account for out of plane scatter as in the cone-beam geometry[96].

4.3. Doses and Imaging Protocols

When quantifying imaging doses the convention is to quote a single numerical value indicative of the absorbed dose (either in the patient, a phantom or the detector) for a particular radiographic projection or CT slice. Alternatively, a single metric may be produced which represents the dose burden from the sequence of images associated with a particular study (e.g. a standard CT scan consisting of a stack of slices, or a fluoroscopy cine sequence). Furthermore, it is desirable to be able to convert such a metric to effective dose in order to evaluate the risk associated with an exposure and aid in optimisation or justification exercises.

Ideally, the quantity chosen to represent CT doses would be simply the absorbed dose to a particular point. However, this is problematic because dose deposition is uniform neither across the plane of the CT slice nor the width of the slice or volume irradiated. In addition, because the couch advances during conventional axial or helical scanning there is the question of where to place the measuring instrument, or indeed how large it should be.

Dose quantities for CT imaging are based on the concept of the 'CT Dose Index' (CTDI)[147, 156, 377], which is reported as an absorbed dose in mGy. There are a number of related quantities:

- **CT Dose Index (CTDI).** The absorbed dose resulting from a single rotation, integrated over a sufficiently long length to cover the entire dose profile of the slice or volume irradiated, including any tails, then divided by the nominal thickness of imaged volume.

- **$\mathbf{CTDI_{air}}$.** CTDI in air, measured with the ionisation chamber in air. (i.e. no scatter conditions.)

- **$\mathbf{CTDI_{100}}$.** CTDI in air, but measured in a cylindrical PMMA phantom and integrated over a length of 10 cm. Two standard phantom are utilised, both of length 14 cm: one 16 cm diameter for head examinations, another 32 cm diameter for body studies.

- **$\mathbf{CTDI_w}$.** Weighted average of $CTDI_{100}$ across the scan plane. Measurements are usually performed at the centre of the phantom and at cardinal compass points 1 cm from the periphery. It is calculated using the following formula:

$$CTDI_w = \frac{1}{3}CTDI_{100}\,(\text{centre of phantom}) + \frac{2}{3}\overline{CTDI_{100}\,(\text{periphery})} \qquad (4.1)$$

- **$CTDI_{vol}$.** Same as $CTDI_w$ but taking into account the effect of couch motion during a helical scan.

$$CTDI_{vol} = CTDI_w / \text{pitch} \tag{4.2}$$

where pitch is the ratio between the distance advanced per tube rotation and the nominal slice thickness (e.g. a couch advance of 1.5 cm / rotation and a slice thickness of 1.0 cm would be a pitch of 1.5.) $CTDI_{vol}$ is considered most representative of the patient dose burden from a single CT slice.

- **Dose Length Product (DLP)**. This is indicative of the dose burden from a whole CT study. It is calculated using

$$DLP = CTDI_{vol} \times \text{scan length} \tag{4.3}$$

and is reported in units of mGy-cm. Calculation software exists to convert DLP to effective dose for different regions of the body, given technical information about the scanner involved and the length of body irradiated[150].

Established and preferred practice in diagnostic imaging is to achieve the 10 cm integration length for $CTDI_w$ by performing measurements using a 10 cm long pencil ionisation chamber. As long as this completely encompasses the dose profile of the slice or volume irradiated then the full contribution from the exposure is integrated over the length of the chamber. However, whilst this approach is straightforward to apply in conventional CT it is more difficult to extend to CBCT. As discussed in section 4.2 the irradiated volume of Acuity or OBI scans may be of the order of 14–17 cm long, thereby extending beyond the end of the chamber. Measurements performed as part of this work, and those of other researchers[8, 201, 313], suggest this may result in an overestimation in dose measurement of the order of 10% due to scatter from the volume irradiated beyond the chamber contributing to the measured signal. An alternative, and theoretically equivalent, approach is to use a small volume ionisation chamber (such as a 0.6 cc Farmer chamber) and measure the dose at the centre of an irradiated volume of length 10 cm. This is potentially more robust for CBCT measurements and is favoured by the radiotherapy community because of the widespread experience of utilising small volume instruments for similar measurements. However, this approach is also not ideal because it means forcing CT equipment to irradiate a 10 cm volume which may not be representative of the collimation and thus scatter conditions of real patient scan protocols. At present there is considerable debate in the imaging community over which methodology should be applied to CTDI assessments,

although current guidance is leaning towards the use of the small volume chamber[250, 276].

Another caveat is that UK convention (as described above) is to report $CTDI_w$ as a dose **to air**, measured in a PMMA phantom, whereas the convention in the USA is to report dose **to PMMA**, measured in a PMMA phantom. This can lead to measurement errors of the order of 10%[377] (the reported dose to PMMA being 10% less than the dose to air) and cause considerable confusion, especially when comparing measurements against manufacturers' specifications or against results in the literature. However, experience has indicated that regardless of the final approach it is more important to be consistent when making measurements between modalities if the goal is to perform dose intercomparisons.

In this current work CT doses are reported following the UK convention: in terms of $CTDI_{vol}$ as a dose to air.

Both conventional CT scanners allow the user to initiate scans according to predefined protocols or to programme the technique setting manually. A wide range of settings can be adjusted, including kVp, tube current (mA), rotation time, pitch, scan and display fields of view and reconstruction filter settings. However, in both Edinburgh and Oxford the tendency was for radiographers to pre-programme a limited number of protocols appropriate for the majority of patients. If required, these could be adjusted prior to scanning for any particular patient (such as increasing the tube current for a large patient). Therefore, only a limited number of technique settings were considered in this study, with these being chosen to be representative of those employed for real patient scans. CT scan protocols used to acquire the images discussed throughout the rest of this chapter are listed in table 4.2 for the HiSpeed FX/i system in Edinburgh and table 4.3 for the HiSpeed QX/i in Oxford.

Similar to the conventional CT scanners, it is possible to perform CBCT scans using the OBI and Acuity systems using a whole range of different acquisition settings. However, full flexibility is only provided in the equipment's service mode, with the clinical interface being entirely protocol driven. This means that the radiography staff are presented with a limited number of scanning modes from which they can choose, with the only modifications possible at scan-time being:

- **Axial FOV.** Although this is restricted to $\leq$ 25.5 cm full-fan and $\leq$ 45.0 cm but $>$ 25.5 cm half-fan.

- **Longitudinal FOV.** Up to 17 cm full-fan and 16 cm half-fan.

- **Reconstructed Slice Thickness.** From 1.0 mm to 5.0 mm, in steps of 0.5 mm, and 10.0 mm.

- **Matrix size of reconstructed image.** 128×128, 256×256, 384×384 or 512×512. It is suggested that smaller matrix sizes might be used to reduce image reconstruction times, although these were rarely found to be problematic. Ideally, the same matrix size would be used for all imaging scenarios. However, although it was expected that the larger pixel size from smaller image matrices would result in a degradation of modulation transfer function (MTF), it was also thought that there might be a corresponding improvement in noise characteristics and possibly a boosting of contrast-to-noise ratio (CNR). Detailed performance comparisons of the 384×384 and 512×512 options were therefore undertaken to investigate whether the differences between them might be clinically relevant.

Default OBI and Acuity scan protocols are listed in tables 4.4 and 4.5 respectively. All of these were considered during this study.

Index	Type	Pitch	kVp	Slice (mm)	Current (mA)	Rot time (s)	SFOV (cm)	DFOV (cm)	$CTDI_{vol}$ (mGy)
E1	Axial	N/A	80	3.0	150	1.0	50	50	4.8
E2	Axial	N/A	80	3.0	300	1.0	50	50	9.7
E3	Helical	1.0	80	3.0	200	1.0	50	50	6.5
E4	Axial	N/A	120	3.0	150	1.0	50	50	12.4
E5	Axial	N/A	120	3.0	300	1.0	50	50	24.8
E6	Helical	1.0	120	3.0	200	1.0	50	50	16.6
E7	Axial	N/A	140	3.0	150	1.0	50	50	17.2
E8	Axial	N/A	140	3.0	300	1.0	50	50	34.3
E9	Helical	1.0	140	3.0	200	1.0	50	50	22.9
E10	Helical	1.0	120	5.0	200	1.0	50	50	16.6
E11	Helical	1.0	140	5.0	200	1.0	50	50	22.9
E12	Axial	N/A	80	5.0	200	1.5	25	25	22.7
E13	Axial	N/A	120	5.0	200	1.5	25	25	55.6
E14	Axial	N/A	140	5.0	200	1.5	25	25	74.7

Table 4.2. – Edinburgh CT scan protocols considered in this work. All scans were reconstructed using a 'Standard' algorithm and convolution kernel.

Index	Type	Pitch	kVp	Slice (mm)	Current (mA)	Rot time (s)	SFOV (cm)	DFOV (cm)	$CTDI_{vol}$ (mGy)
O1	Axial	N/A	120	5.0	200	1.0	50	50	25.5
O2	Axial	N/A	120	5.0	200	1.0	50	65	25.5
O3	Axial	N/A	80	2.5	200	1.0	25	25	18.8
O4	Axial	N/A	100	2.5	200	1.0	25	25	34.2
O5	Axial	N/A	120	2.5	200	1.0	25	25	52.4
O6	Axial	N/A	120	2.5	200	1.0	50	50	25.5
O7	Axial	N/A	120	2.5	200	1.0	50	65	25.5
O8	Axial	N/A	140	2.5	200	1.0	25	25	72.8

Table 4.3. – Oxford CT scan protocols considered in this work. All scans were reconstructed using a 'Standard' algorithm and convolution kernel.

Abbrev	Mode Name	Acq Angle (°)	Fan Type	kVp	mAs	$CTDI_{vol}$ (mGy)
SD Head	Standard-dose Head	200	Full	100	145	4.3
LD Head	Low-dose Head	200	Full	100	72	2.2
HQ Head	High-quality Head	200	Full	100	720	21.6
Pelvis	Pelvis	360	Half	125	680	19.7
PSL	Pelvis Spot-light	200	Full	125	720	16.0
Thorax	Low-dose Thorax	360	Half	110	262	5.2

Table 4.4. – Default Varian OBI CBCT scan protocols. Source-isocentre-distance (SID) = 100 cm, source-detector-distance (SDD) = 150 cm and nominal slice thickness was 2.5 mm for all protocols.

Abbrev	Mode Name	Acq Angle (°)	SDD (cm)	kVp	mAs	$CTDI_{vol}$ (mGy)
MS 150	Multiscan 150	360	150	125	780	20.5
MS 160	Multiscan 160	360	160	125	780	19.3
SD 150	Standard Dose 150	360	150	125	1300	35.4
SD 160	Standard Dose 160	360	160	125	1300	32.8

Table 4.5. – Default Varian Acuity CBCT scan protocols. All modes are automatically half fan if the field of view > 26 cm and full fan if < 26 cm. SDD = source-detector-distance. Nominal slice thickness was 2.5 mm for all protocols.

4.4. HU to Electron Density Conversion Curves

Dose calculations for treatment planning may utilise relative electron density information to account for inhomogeneities in the patient[145, 242]. Mapping of the pixel values in CT images to the electron densities of the tissues they represent allows these corrections to be performed on a pixel by pixel basis, thus potentially increasing the accuracy of the dose calculation[144, 265].

Pixel values of CT images are calibrated in terms of Hounsfield Units (HU)[168]. These are defined as

$$HU = 1000\left(\frac{\mu - \mu_w}{\mu_w}\right) \tag{4.4}$$

where μ is the linear attenuation coefficient of the material in the pixel and μ_w is that of pure water. μ depends upon atomic number, electron density and X-ray beam spectrum. A complex relationship therefore exists between HU, material properties, the physics of the X-ray tube and the accelerating potential (kVp) used to generate the beam. However, for materials with similar interaction properties (i.e. effectively μ proportional to electron density ρ_e) then conversion curves between HU and the ED can be determined for particular beam spectra.

ICRU Report 42[144] defines 'standard' conversion curves for tissues in the body:

$$\rho'_e\,(soft) = 1.000 + 0.001HU \quad HU \leq 100 \tag{4.5}$$

$$\rho'_e\,(bone) = 1.052 + 0.00048HU \quad HU > 100 \tag{4.6}$$

where ρ'_e is the electron density relative to that of water ($\rho'_{e,water} \equiv 1.0$). These are based on work by Battista and Bronskill[18] and were determined by scanning radiologically equivalent tissue-mimicking materials using a Picker Synerview 120 CT scanner (Picker International, Ohio, USA) operating at 135 kVp and with an X-ray beam of effective energy $\tilde{E}$ of 72 keV. In their paper it is suggested that these curves should calculate electron density accurately to within 5% for all beams of effective energy 60 keV < $\tilde{E}$ < 80 keV. With reference to the values in table 4.1, it was therefore expected that the 120 kVp and 140 kVp beams of the HiSpeed FX/i and HiSpeed QX/i scanners and the default 125 kVp beam of the Acuity CBCT system should demonstrate electron density calibration curves consistent with this work.

Curves have also been determined by other workers, either theoretically or

by measurement[55, 186, 228, 265, 314, 336], and although there are significant differences between the various curves they all follow the same general trend, comprising two linear segments: one extending from air, through lung and soft tissue, and a second, steeper curve from $\sim\rho'_e > 1.1$ for bone. The bone segment is steeper because the increased photoelectric interactions due to the K-edge of calcium causes μ and thus the measured HU to increase more rapidly with ρ'_e.

It is attractive to be able to use the ICRU 42 curve if possible for a number of reasons. Firstly, it is the default curve in many treatment planning systems including Varian's Eclipse, which is used in Oxford and Edinburgh, so that by employing this default curve the underlying calculation system will be configured similarly to others worldwide. Also, this approach allows comparisons against other treatment planning systems which are configured with the same standard curve. Finally, the curve increases monotonically and is continuous in both electron density and HU, whereas some of the others have a discontinuity at the interface between bone and soft tissue. Although these may be better representations of reality, the well-behaved ICRU 42 curve provides an unambiguous one to one mapping between electron density and HU, making calculations using either quantity more straightforward.

A complicating factor when converting CT pixel values to electron densities is that the CT reconstruction algorithms are sensitive to the presence of very attenuating materials, with the result being that high-density objects in the field may influence the pixel values of other structures. Sometimes this effect may be asserted over the entire length of the field of view so that a high density object at one edge of the scan may affect the pixel values of another object at the other. Therefore, once a CT calibration curve has been established it is important to verify it remains true under typical patient imaging scenarios. However, to do this it is necessary to understand by how much the pixel values of an object may be permitted to deviate from the ideal curve.

IPEM Report 81[154] recommends that electron densities reported for CT images in the treatment planning system are within ±1% of the real values for water, and ±2% for bone and lung. However, these are very tight tolerances and cannot generally be achieved in practice due to the wide range of imaging scenarios typically encountered during radiotherapy CT scanning. Instead, these may be more appropriately applied in a routine QA programme as tolerances about measured baseline values outside which CT scanner performance must not drift.

More generally, there is consensus that radiotherapy dose delivery should

be accurate to within ±3% (1 SD) at the specification point and ±5% (1 SD) elsewhere in the target volume[154]. A pragmatic limit for the contribution to this total uncertainty by the treatment planning system calculation may therefore be ±2% (1 SD). As a result, an appropriate absolute limit on the change in dose resulting from fluctuations of the HU-electron density conversion curve might be set at ±2% (which is effectively 3 SD)[178].

Differentiating equations 4.5 and 4.6 yields a relationship between how a change in HU dHU will cause a change in electron density $d\rho'_e$.

$$d\rho'_e\,(soft) = 0.001 \quad dHU \leq 100 \tag{4.7}$$

$$d\rho'_e\,(bone) = 0.00048 \quad dHU > 100 \tag{4.8}$$

Kilby *et al.*[178] model treatment by a 6 MV photon beam of a thick tissue section containing 20 cm water, 10 cm lung and 7 cm bone. This is a worst-case scenario in that the dose calculation is very sensitive to changes in HU. Their study indicates that a difference in dose of 2% would be caused by changes in ρ'_e of ±0.03 for water, ±0.05 for lung and ±0.08 for bone. According to equations 4.7 and 4.8 the allowable changes in HU are therefore ±30 HU for water, ±50 HU for lung and ±167 HU for bone. These are similar to the values arrived at by Kirwin *et al.*[184], whose work uses Kilby *et al*'s tolerances but with calibration curves based on Thomas[336] and Knoos *et al.*[186] rather than that of ICRU 42.

In this current work, the limits calculated from Kilby *et al.*'s study are applied around the ICRU 42 curve to evaluate whether a particular measured CT calibration curve is within tolerance. It is accepted that tighter tolerances may be required for electron planning or in more critical dose calculation scenarios[145, 178].

4.5. RMI-467 Electron Density Phantom

One test object for characterising the HU to electron density conversion curve is the RMI-467 phantom (Gammex, Wisconsin, USA) which was developed by Constantinou and Harrington[55]. This phantom comprises a 33 cm diameter disc containing 16 holes of diameter 25 mm in which tissue-equivalent materials with known electron densities can be inserted. Rods of 12 different materials are included with the phantom, along with a container that can be filled with pure water. Because there is some variability in the manufacturing process each purchased phantom is supplied with a table of measured electron density values

for the inserts included with that phantom. In addition, the phantom contains a number of sub-millimetre holes which can be used for distance measurements. A photograph and schematic diagram of the RMI-467 phantom containing the electron density inserts in their standard configuration is shown in 4.1. This figure also includes a table of the insert electron densities for the phantoms available in Edinburgh and Oxford.

An IQWorks analysis tree was constructed for the RMI-467 phantom and included:

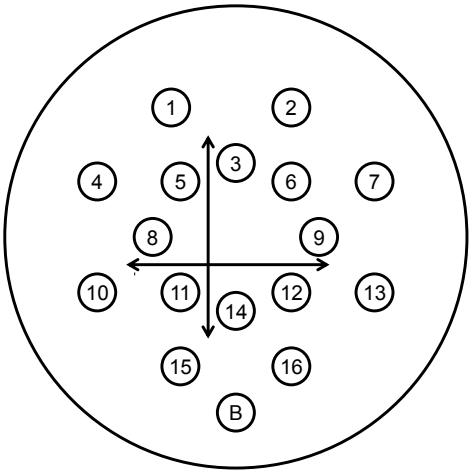
- Identification of the outer edge of the phantom, including calculating of phantom width, height and the centre of extremes of the edge. This was used to localise other ROIs. However, its applicability to the assessment of geometric linearity was limited due to difficulties with mounting the phantom on the flat radiotherapy couch, and it not being possible to lower the couch of the Oxford scanner sufficiently to centre the phantom in the FOV.

- Calculation of the mean and standard deviation in ROIs placed on each electron density insert. These results are plotted as a function of electron density.

- Measurement of the horizontal and vertical distances between two pairs of holes, taking the centre of each hole as the location of the minimum pixel value within the search ROI. These measurements are indicated by the arrows on the schematic diagram in figure 4.1.

- Calculation of contrast (using equation 2.8) and CNR (using equation 2.5) between all inserts and the phantom base material.

- Generation of a report containing results of the above analyses and verifying all material HU agree with the measured baseline within a tight tolerance of ±10 HU.



No.	Material	Electron Density Rel to Water ρ'_e	
		Edinburgh (SN 802428-1027)	Oxford (SN 802428-1098)
1	CT Solid Water	0.988	0.986
2	LN-300 Lung	0.268	0.262
3	SB3 Cortical Bone	1.693	1.695
4	LN-450 Lung	0.436	0.459
5	BR-12 Breast	0.957	0.958
6	BRN-SR2 Brain	1.047	1.047
7	CT Solid Water	0.988	0.986
8	True Water	1.000	1.000
9	AP6 Adipose	0.937	0.930
10	B200 Bone Mineral	1.105	1.111
11	CB2-30% CaCO3	1.275	1.273
12	CB2-50% CaCO3	1.470	1.466
13	IB Inner Bone	1.105	1.097
14	CT Solid Water	0.988	0.986
15	LV1 Liver	1.077	1.072
16	CT Solid Water	0.988	0.986
B	Phantom Body	0.989	0.986

Figure 4.1. – RMI-467 phantom with electron density inserts in standard configuration[55]. Top Left–Photograph of the phantom aligned prior to scanning. Top Right–Schematic diagram of the phantom. Bottom–Table of materials and electron densities for the individual phantoms used in Edinburgh and Oxford.

A series of scans was acquired using the HiSpeed FX/i scanner in Edinburgh, using protocols E1–E9 in table 4.2. Different combinations of X-ray tube accelerating potential (kVp), tube current (mA) and whether helical or axial mode were employed were chosen in order to investigate their influence on the electron density calibration curve. The images were processed by IQWorks and the measured electron density curves compared against the ICRU 42 reference curve using the tolerances described in section 4.4 to verify accurate dose calculation to within ±2%.

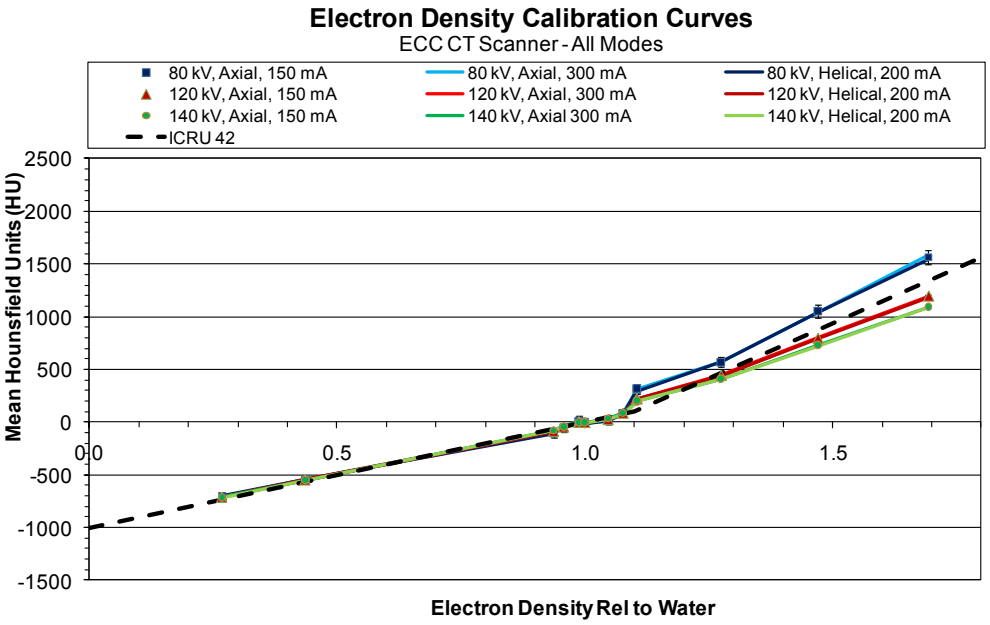
From the plots in figure 4.2 it is clear that whilst kVp has a considerable effect on the electron density curve, the influences of tube current, and whether axial or helical mode is employed, are very small. Furthermore, no measured curve was in perfect agreement with the ICRU 42 reference, with all curves following similar trends and moving either side of the reference curve across the electron density range. Only the 120 kVp curve is within the specified tolerance band at all electron densities, with the 140 kVp curve falling outside tolerance for the highest density points and the 80 kVp curve outside tolerance at multiple points across the range. These results suggest that whilst axial and helical modes, as well as the X-ray tube current, can be set as appropriate for the clinical application, care needs to be taken over the choice of kVp, especially if the ICRU 42 reference curve is to be employed in the treatment planning system. It was expected that the 120 kVp curve would agree well with the ICRU 42 curve because the effective energy of the beam spectrum (60 keV) lies at the lower limit of the 60 keV – 80 keV range determined by Battista and Bronskill[18]. However, it was somewhat surprising that the 140 keV curve had poor agreement for the highest density inserts because its effective energy (65 keV) is closer to the centre of the ideal range. This may be due to the influence of beam hardening artefacts, which were clearly visible in all images, the location of the high attenuation inserts within the FOV, or related to the detail of how the scanner is intrinsically calibrated. There also appears to be a trend in deviation from the ICRU curve for bony materials ($\rho'_e > 1.2$): at 80 kVp HU is grossly overestimated ($\gtrsim +2\%$), at 120 kVp it is slightly underestimated ($\gtrsim -2\%$) whilst at 140 kVp it is further underestimated ($\lesssim -2\%$). These results suggest that the optimum acceleration potential is somewhere around 100 kVp.

Because of the advantages associated with utilising the ICRU 42 curve it was decided to perform all radiotherapy CT scans in Edinburgh at 120 kVp.

In Oxford Cancer Centre a GE HiSpeed QX/i CT scanner is dedicated to radiotherapy applications. As described in section 4.2, this is similar to the

Edinburgh scanner except that it is a multi-slice unit, capable of acquiring 4 slices simultaneously, and can also scan at 100 kVp. In addition, it incorporates a 'wide FOV' facility which utilises interpolation algorithms to extend the standard 50 cm FOV over which projection data are acquired to 65 cm. This is advantageous for the treatment planning of large patients, or for treatment sites such as breast where the patient set-up and immobilisation devices cause the volume of interest to be near the edge of the scanner bore. However, there was concern that the electron density values may be adversely affected by the extrapolation algorithms.

Scans of the RMI-467 phantom were acquired to verify the electron density curve for the wide FOV mode. The set-up was the same as that described above and utilised scan protocols O1 and O2 in table 4.3, which are the same except one has SFOV and DFOV = 50 cm, whilst the other has SFOV = 50 cm and DFOV = 65 cm. Only the 120 kVp beam was considered because the decision had already been taken in Oxford to scan all radiotherapy patients at 120 kVp. Figure 4.3 contains the curves generated by IQWorks and demonstrates that there is negligible difference between the two modes. Furthermore, it is reassuring that the Oxford curves are nearly identical to the Edinburgh ones, suggesting that beam spectra and calibration processes are similar between the two models of GE scanner. From informal discussions this does indeed seem to be the general consensus in the community and should be expected because of the similar total filtration of the X-ray tubes employed (see table 4.1).



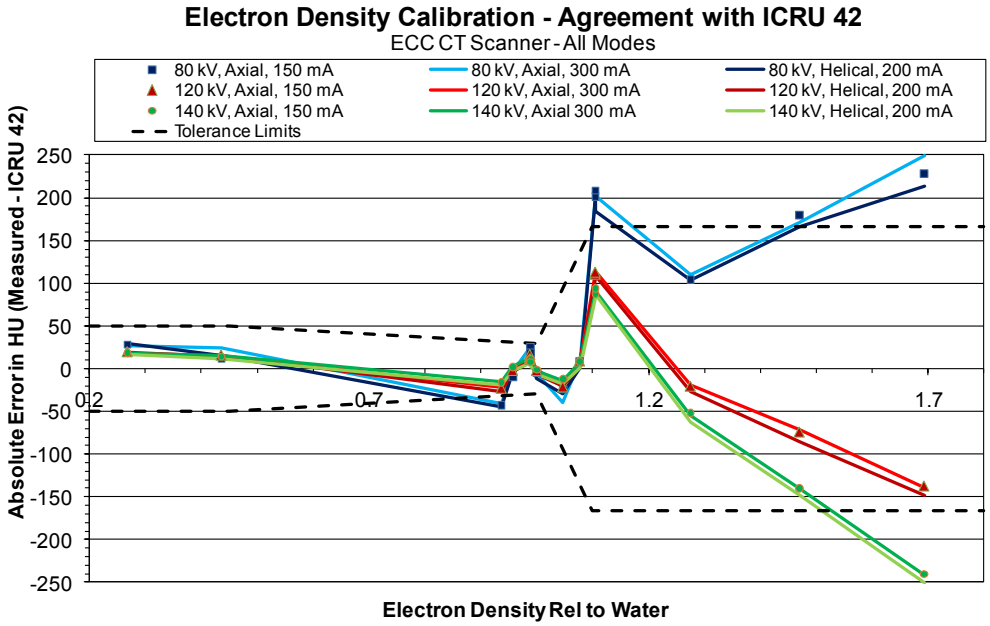


Figure 4.2. – Electron density calibration curves of the radiotherapy CT scanner in Edinburgh Cancer Centre (ECC), measured using the RMI-467 phantom and IQWorks. Error bars indicate ±1 SD in the pixel values in each region of interest. Top–Measured calibration curves, with ICRU-42 curve present as the reference ideal curve. Bottom–Deviation of measured curves from ICRU-42 curve, with the tolerance limits being those required to maintain dose calculation accuracy to within ±2%.

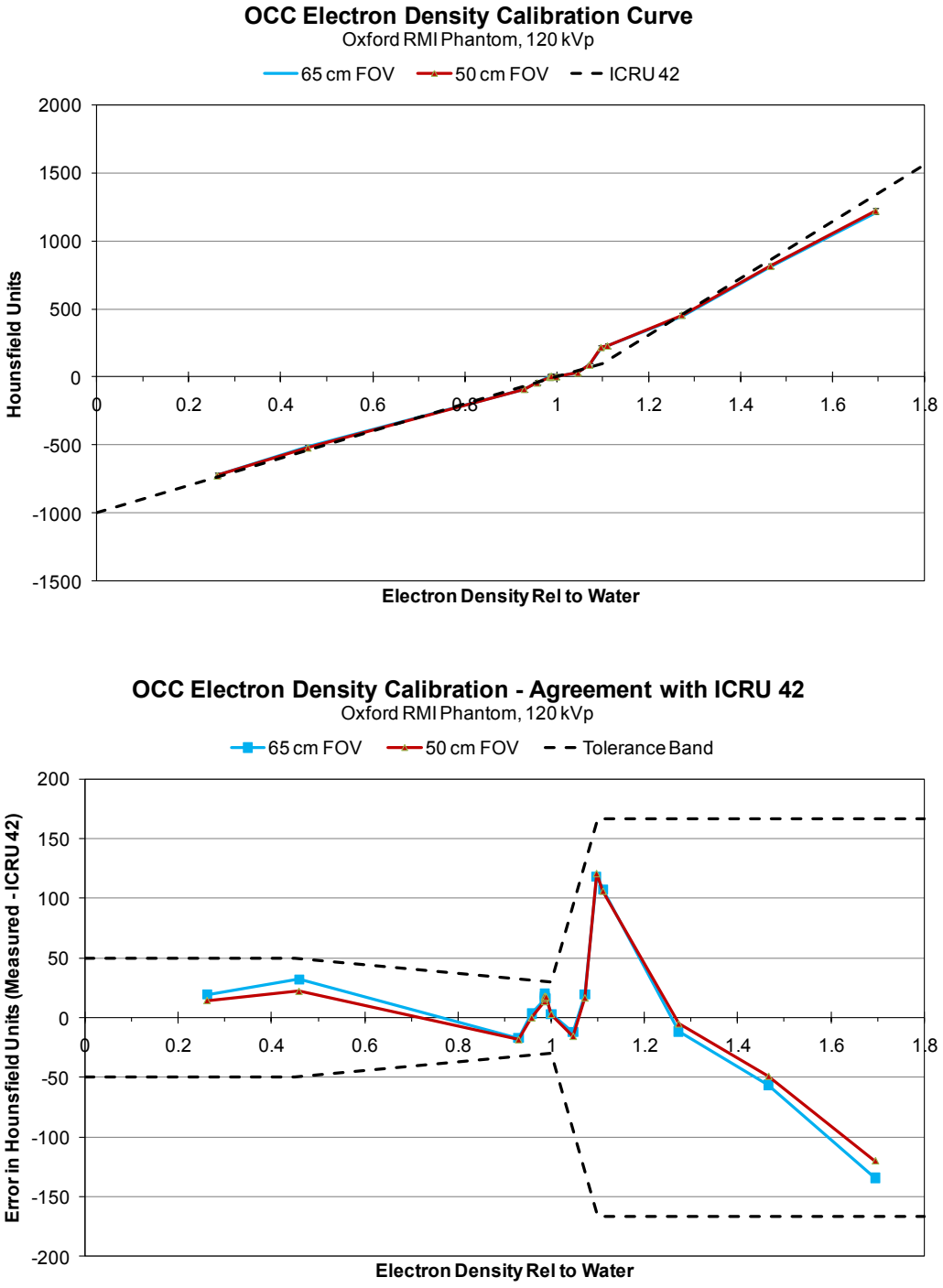


Figure 4.3. – Electron density calibration curves of clinical imaging modes of the radiotherapy CT scanner in Oxford Cancer Centre (OCC), measured using the RMI-467 phantom and IQWorks. Top–Measured calibration curves, with ICRU-42 curve present as the reference ideal curve. Bottom–Deviation of measured curves from ICRU-42 curve, with the tolerance limits being those required to maintain dose calculation accuracy to within ±2%.

In addition to being used for detailed performance characterisation the RMI-467 phantom is scanned monthly in Oxford as part of the routine CT QA programme. Running the IQWorks analysis tree on historical images enabled a database of results to be constructed from which time-trends could be extracted. Four reference materials were selected for detailed historical analysis: lung, water, brain and cortical bone. These were chosen as a representative sample from across the electron density range. Although scan protocols O1 and O2 in table 4.3 were both utilised routinely for the QA measurements, only the results of O1 (the 50 cm DFOV protocol) are presented here. This is because consistency in trends was observed between both protocols and because the 65 cm DFOV protocol was only introduced recently – within the last year – whereas the 50 cm protocol has been in regular use over the 6 year lifespan of the scanner.

Mean pixel value in each of the reference inserts is plotted as a function of time in figure 4.4. It is evident that the CT scanner performance is remarkably stable, with all materials lying well within the tolerances required to prevent a 2% change in dose calculation. Furthermore, for all materials other than bone an absolute limit of ±5 HU would be an appropriate tolerance for routine QA. Although bone exhibits significantly greater variability it is still within ±12 HU across the whole time-span. Indeed, given the discussion above about CT performance being very sensitive to the positioning of highly attenuating objects, one might expect that small alignment changes, whether the phantom is mounted using a particular jig and even the longitudinal position of the couch may all have an impact on this.

Contrast-to-noise ratio (CNR) for each reference material, relative to its baseline value, is plotted over time in figure 4.5. Some interesting conclusions can be drawn from this. When the CT scanner is operating over a period of stability the CNR for all materials lies within a well defined tolerance band of approximately ±10%. However, the CNR is sufficiently sensitive to highlight a change in performance due to other objects lying within the FOV, with the 'Bad Couch Position' point indicating when the phantom was positioned near the very end of the flat couch-top so that reinforcement structures appeared in the CT scan. Although visually these did not appear to have a major impact it is clear that they resulted in an increase in noise across the FOV, presumably because of the greater attenuation caused by them. Furthermore, definite changes in CNR following software upgrades are also visible, with the first upgrade indicated boosting CNR by approximately 20% through reducing image noise. Interestingly, there may be correlation between the dates of the two software

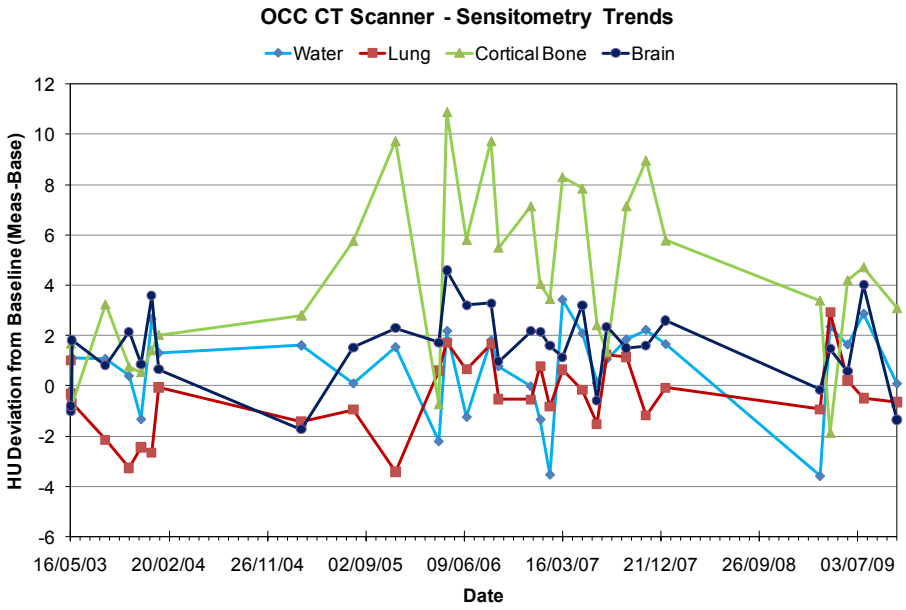


Figure 4.4. – Sensitometry time-trends of the radiotherapy CT scanner in Oxford Cancer Centre (OCC), measured using the RMI-467 phantom. Data are presented for four reference materials all scanned at 120 kVp.

upgrades and the period of particular instability of the bone HU (shown in figure 4.4), although there is insufficient data for this to be conclusive.

Time-trends of the distances measured between the holes in the phantom are presented in figure 4.6, from which it is clear that the geometric linearity of the scanner is very stable, never deviating by more than ±0.5%. Points in the figure appear to fall into discrete levels due to the step change observed if moving a distance of one whole pixel (0.98 mm at a 50 cm FOV).

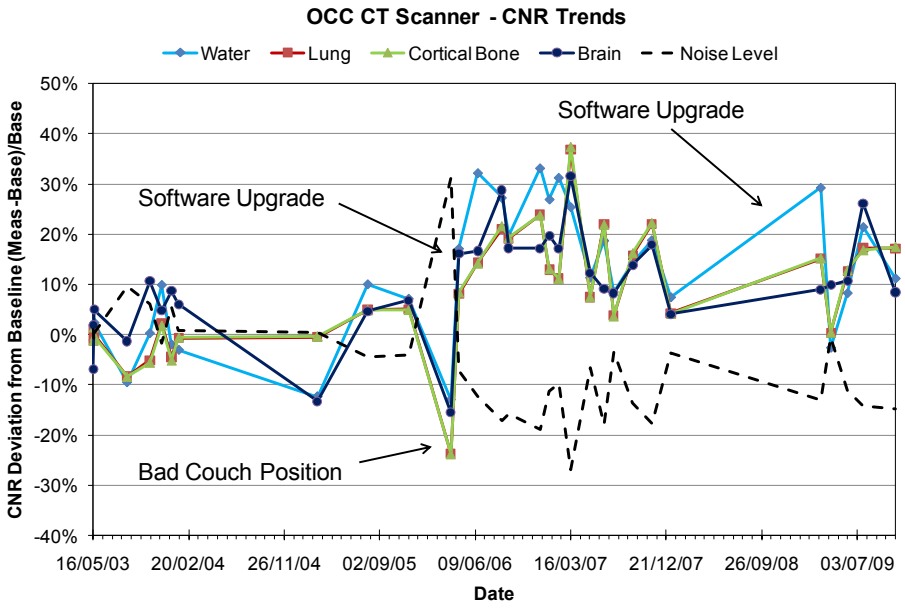


Figure 4.5. – CNR time-trends of the radiotherapy CT scanner in Oxford Cancer Centre (OCC). Data are presented for four reference materials all scanned at 120 kVp. Changes in performance due to software upgrades and couch artefacts are clearly visible.

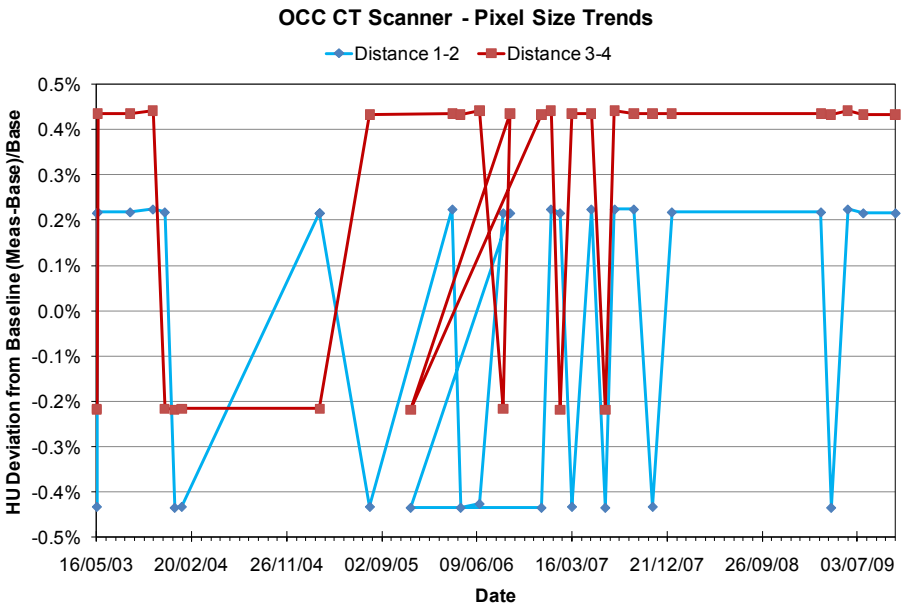


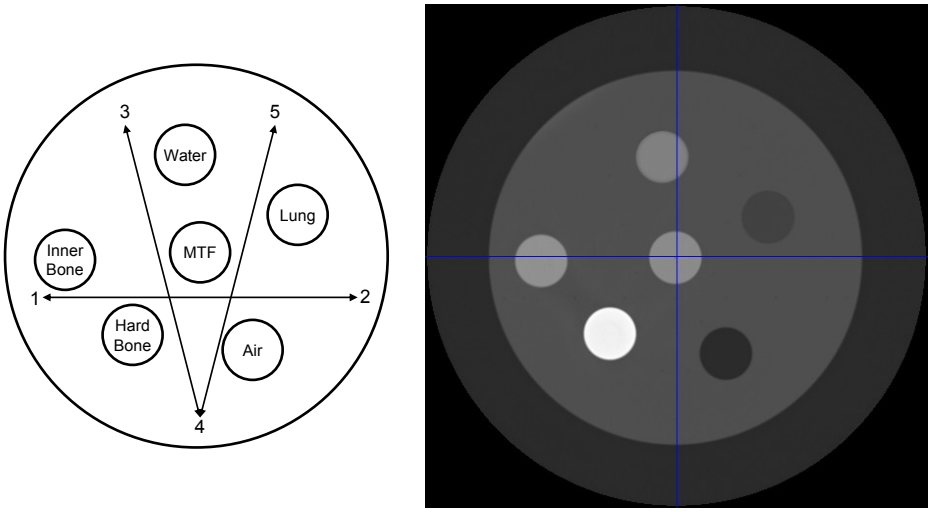
Figure 4.6. – Pixel size time-trends of the radiotherapy CT scanner in Oxford Cancer Centre (OCC), measured using the RMI-467 phantom.

4.6. Varian CT Performance Phantom

Another phantom suitable for verifying the HU – electron density conversion curve is the Varian CT Performance Phantom (Varian Medical Systems, Palo Alto). Although originally intended as a QA and calibration test object for the single-slice CT option of the Varian Ximatron simulator[348], this phantom is nevertheless applicable to other CT systems. Indeed, because the RMI-467 phantom was available only occasionally on loan from the Beatson Oncology Centre, the Varian Performance Phantom formed the basis of the CT QA programme in Edinburgh.

Similar to the RMI-467, the Varian Performance Phantom is a disc containing a number of large holes into which different inserts can be placed, along with a series of sub-millimetre bore holes for distance measurements. However, in addition to electron-density inserts this phantom is supplied with further ones specifically for image quality evaluation. These include a contrast-detail test pattern, angled ramps to assess slice thickness and a 1 mm diameter hole drilled through a PMMA plug to provide an impulse object for modulation transfer function (MTF) evaluation. Although the phantom is less suitable for detailed characterisation of the electron-density conversion curve because it can only accommodate 6 materials simultaneously, as opposed to the 16 of the RMI-467, the additional image quality inserts make it more suitable as a comprehensive test tool.

A CT scan and schematic diagram of the Varian Performance Phantom in the standard configuration employed in Edinburgh is shown in figure 4.7, along with a table of calibrated electron-densities for the material inserts. The IQWorks analysis tree developed for this phantom was similar to that for the RMI-467 phantom, but with the addition of the 'Impulse MTF' module to analyse the MTF hole. The algorithms underpinning the analysis of this module are described in section B.17.4 and are based on the assumption that the point spread function (PSF) is radially symmetrical. As a result, the PSF measured is an average of the true 2D PSF in all directions and the MTF also exhibits radial symmetry in frequency space. This approach is commonly employed in many CT performance phantoms[26, 34, 105, 156]



Insert Material	Electron Density Relative to Water ρ'_e
Water	1.000
Lung	0.244
Inner Bone	1.144
Hard Bone	1.698
Air	0.000
Base (Urethane)	0.456

Figure 4.7. – Varian CT Performance Phantom. Left–Schematic diagram illustrating arrangement of inserts and holes for distance measurements. Right–CT slice through centre of phantom. Bottom–Table of material properties.

To assess the potential of the Varian Performance Phantom for verifying the electron density curve the same experiment was repeated as described above for the RMI-467 phantom, using the same acquisition settings (protocols E1–E9 in table 4.2). From the results in 4.8 it is clear that the 80, 120 and 140 kVp curves follow similar behaviour to those measured with the other phantom: agreement is good for $\rho'_e < 1.1$ but for electron densities above this the 80 kVp beam grossly overestimates HU and is well outside tolerance, the 120 kVp beam slightly underestimates the HU, but is within a tolerance of 1% and the 140 kVp underestimates HU slightly more, but this time is still within the overall tolerance of 2%. These results are reassuring because they essentially validate the two phantoms against each other, indicating that either is a suitable tool for routine checks of electron-density calibration. Possible reasons for the curves being marginally different between the two phantoms could be that the electron density inserts are physically larger in the Varian Performance phantom, compared with the RMI-467 phantom (5 cm diameter as opposed to 2.5 cm) and that the objects are in different geometrical configuration within the FOV. Both of these would result in different beam attenuation conditions for the same projections through the phantom, thus influencing the spectrum of the beam reaching the detector array and affecting the calculation of the 2D linear attenuation map.

As part of the Edinburgh QA programme, images of the Varian Performance Phantom were acquired regularly using protocols E10 and E11 in table 4.3. These were representative of the clinical protocols in use when the scanner was first installed. Both 120 and 140 kVp beams were included in the testing because it was felt that there could be occasions when the improved penetration of the 140 kVp beam could be clinically justified, at the expense of patient dose and a slightly degraded electron density curve. The 80 kVp curve was not considered because this mode was discounted soon after installation as being too far outside the acceptable limits. Initially, the phantom was scanned every fortnight but this frequency gradually decreased to once every six weeks as confidence in the stability of the scanner grew.

A time-trend plot of mean pixel values for the lung, water and hard bone inserts, relative to their baseline values, is presented in figure 4.9. Similar to the Oxford measurements using the RMI-467 phantom the sensitometry of the scanner is very stable, this time with the non-bone inserts falling within ±1.5 HU of the baseline across the whole measurement period. As before, bone also exhibits greater variability, although for the Edinburgh scanner it is at worse -5

HU relative to the baseline. There is clearly asymmetry in these results, with the bone signal tending to be underestimated, but rarely overestimated. It is thought that as the scanner grew older the X-ray tube became less stable, with the highest energy mode being the most difficult to maintain, and that the asymmetry can be attributed to a gradual trend away from baseline. Nevertheless, the absolute change in calibration accuracy is negligible and certainly not clinically significant.

CNR trends are plotted in figure 4.10. These are almost identical to the Oxford results in that, during stable scanner operation, all measurements lie within a band ±10% of the baseline. However, there are no shifts or spikes in performance as were observed for the Oxford scanner. This may partly be due to the alignment of the Varian Performance Phantom being more reproducible: the Varian phantom is hung on a mount over the end of the couch, whereas the RMI-467 is placed in a jig on the couch-top itself.

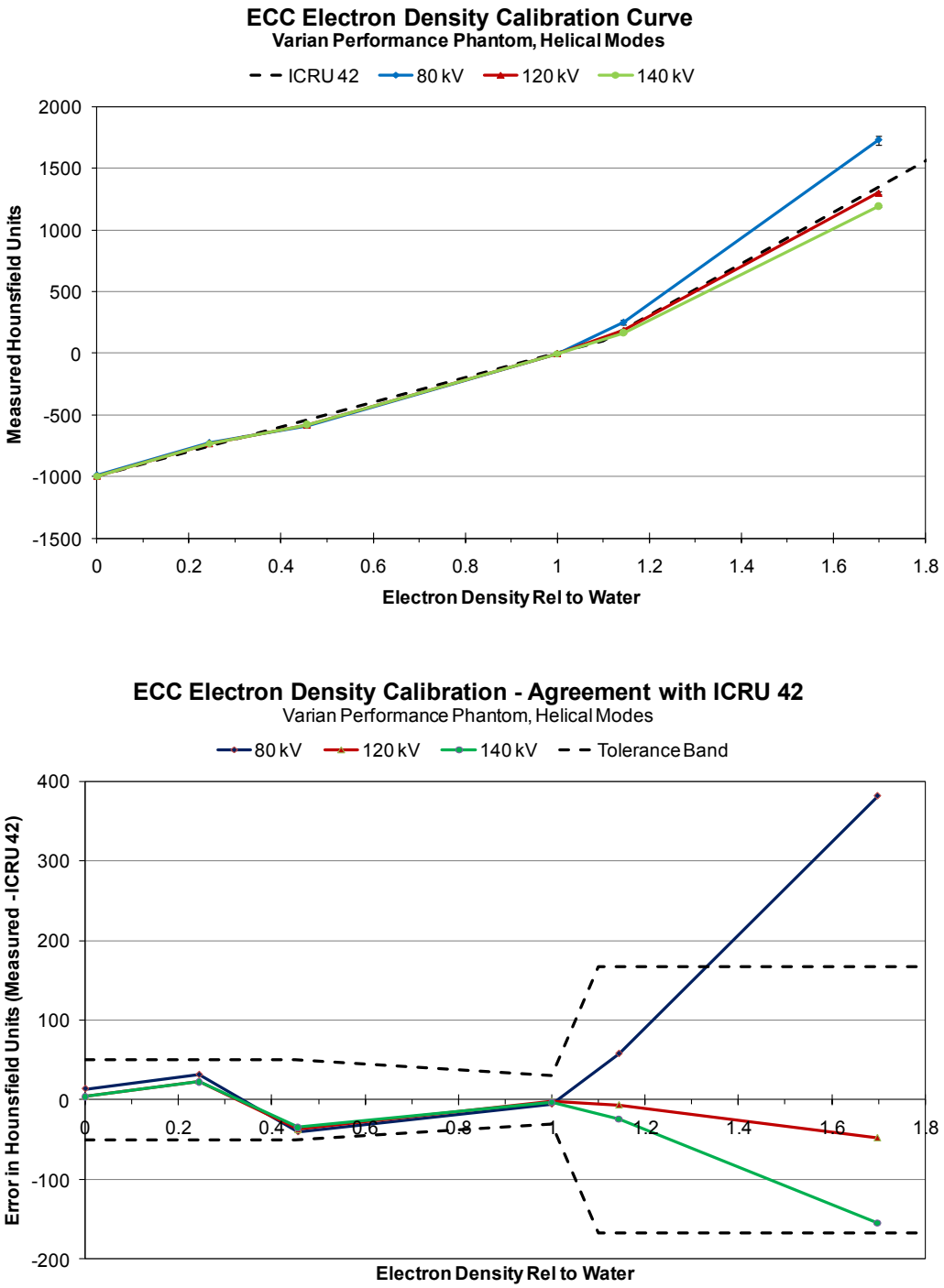


Figure 4.8. – Electron density calibration curves of the radiotherapy CT scanner in Edinburgh Cancer Centre (ECC), measured using the Varian CT Performance Phantom and IQWorks. Error bars are ±1 SD of the pixel values in each region of interest. Top–Measured calibration curves, with ICRU-42 curve present as the reference ideal curve. Bottom–Deviation of measured curves from ICRU-42 curve, with the tolerance limits being those required to maintain dose calculation accuracy to within ±2%.

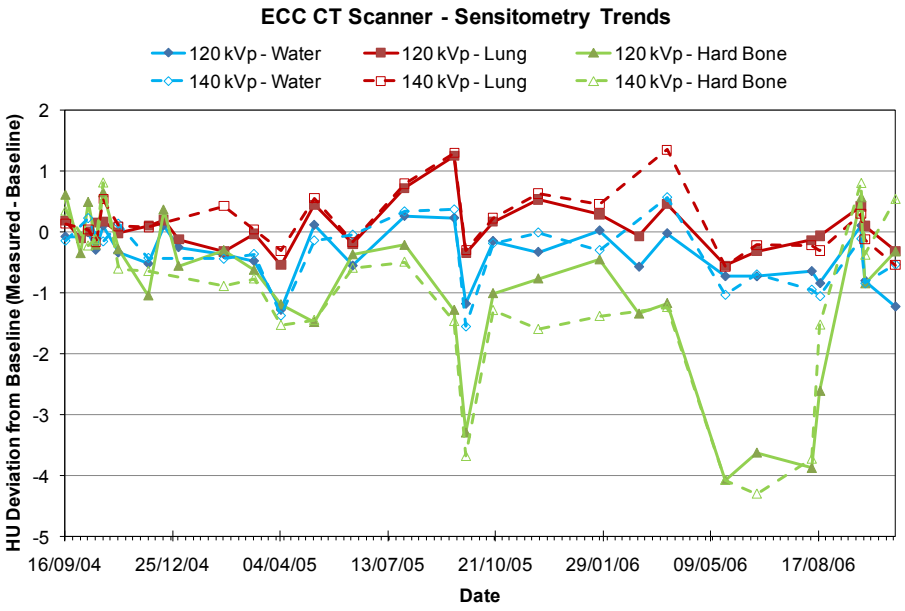


Figure 4.9. – Sensitometry time-trends of the radiotherapy CT scanner in Edinburgh Cancer Centre (ECC). Data are presented for three reference materials scanned at 120 and 140 kVp.

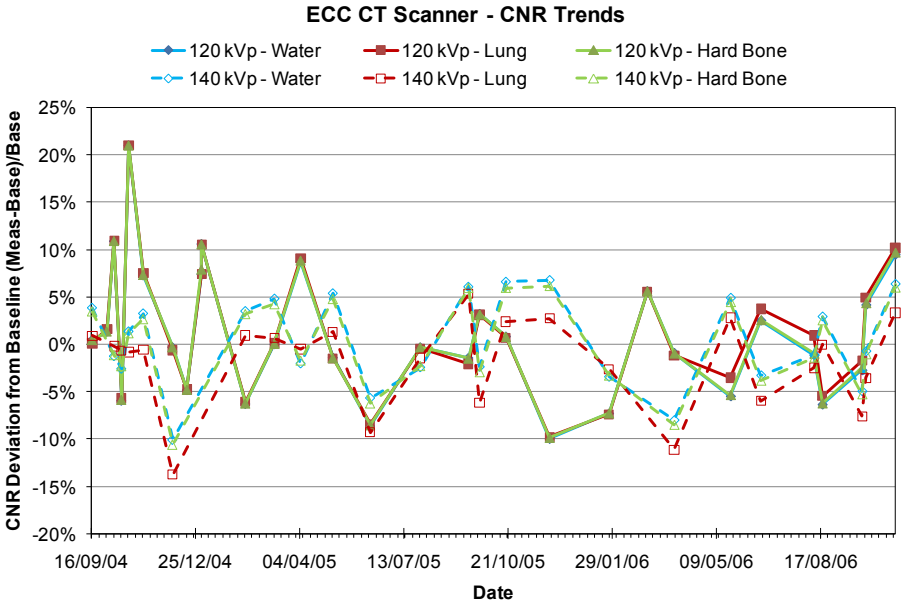


Figure 4.10. – CNR time-trends of the radiotherapy CT scanner in Edinburgh Cancer Centre (ECC). Data are presented for three reference materials scanned at 120 and 140 kVp.

MTF curves calculated from the impulse object at 120 and 140 kVp are shown in figure 4.11, from which it is appears that MTF is insensitive to beam spectrum, with the f_{50} and f_{10} being identical for both curves to 2 significant figures. This is as expected because spatial resolution is known to be influenced by sampling conditions and the nature of the reconstruction algorithm (convolution kernel, pixel matrix, etc.), but should be independent of kVp. As long as the intrinsic contrast between the air filled impulse object and PMMA (which is of the order of 1100 HU for these materials) is sufficient to be discernible above the Poisson noise in the system (which is considerably smaller, having a standard deviation of only 4 HU in the PMMA plug), then the MTF should be similar at different kVps, for the same sampling / reconstruction conditions[84, 169].

Unfortunately, it is difficult to make direct comparisons between these and those of other workers because of subtle differences in acquisition settings. When performing detailed type testing of the HiSpeed FX/i scanner as part of a Medical Devices Agency evaluation exercise[45] the ImPACT team measured $f_{50} = 0.34$ cycles/mm and $f_{10} = 0.59$ cycles / mm for a protocol which was very similar to this one, except that the projection data were acquired using a SFOV of 50 cm but reconstructed at a DFOV of 24.9 cm. The rationale behind this approach is that the sampling conditions may be superior than that afforded by a 512×512 pixel matrix at a full 50 cm FOV, so reconstructing at a small FOV (and thus smaller pixel size) allows the intrinsic spatial resolution of the detector to be accurately assessed. However, it is not necessarily representative of the resolving power of clinical CT scans, which may be limited by the larger presented pixel size to accommodate a larger FOV using the same number of pixels. This is frequently the case in radiotherapy, where large FOV scans are the norm. Nevertheless, the finding that both ImPACT measurements are of the order of 50% greater than those measured in this work is consistent with the reconstructed pixel size being reduced by the same factor between the ImPACT experiment and the one described here. This gives confidence in the results calculated via the Varian Performance Phantom and IQWorks.

From the MTF time-trend analysis in figure 4.12 it is also encouraging that f_{50} and f_{10} lie within ±0.06 cycles / mm of the baseline across the whole time period, again demonstrating the stability of the system.

Geometric linearity was assessed regularly through monitoring both the distances indicated in figure 4.7 and the width and height of the detected phantom edge. Both sets of results, presented in figures 4.13 and 4.14 respectively, indicate a very stable geometry over time. All measurements are within ±0.4%

of baseline, consistent with the results found for the Oxford scanner. It is interesting to note that the phantom width / height results are slightly tighter than those of the hole-based distance measurements. This is because phantom width and height are always measured as a whole number of pixels, whereas any individual distance vector may include components from both matrix directions.

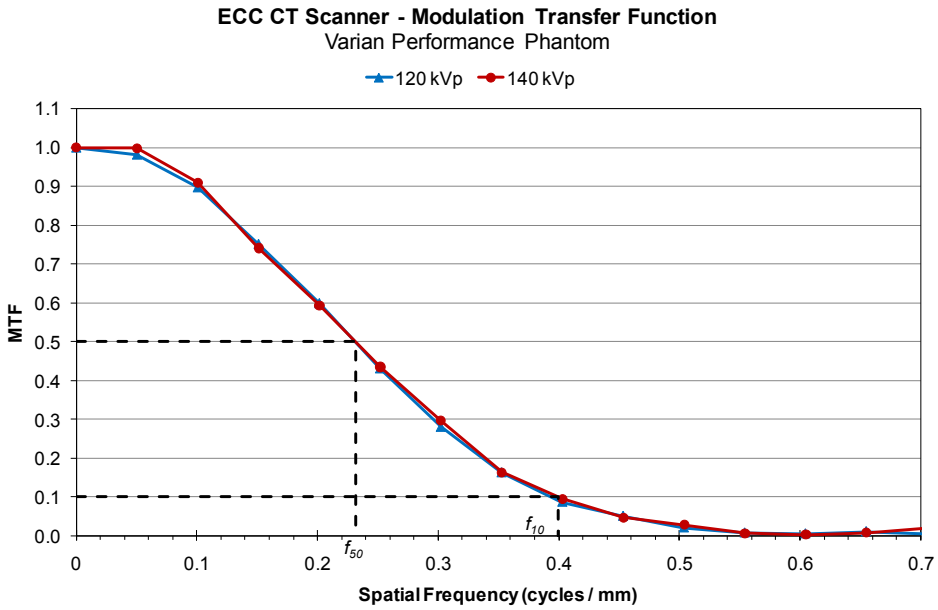


Figure 4.11. – MTF of two clinical imaging modes of the radiotherapy CT scanner in Edinburgh Cancer Centre (ECC).

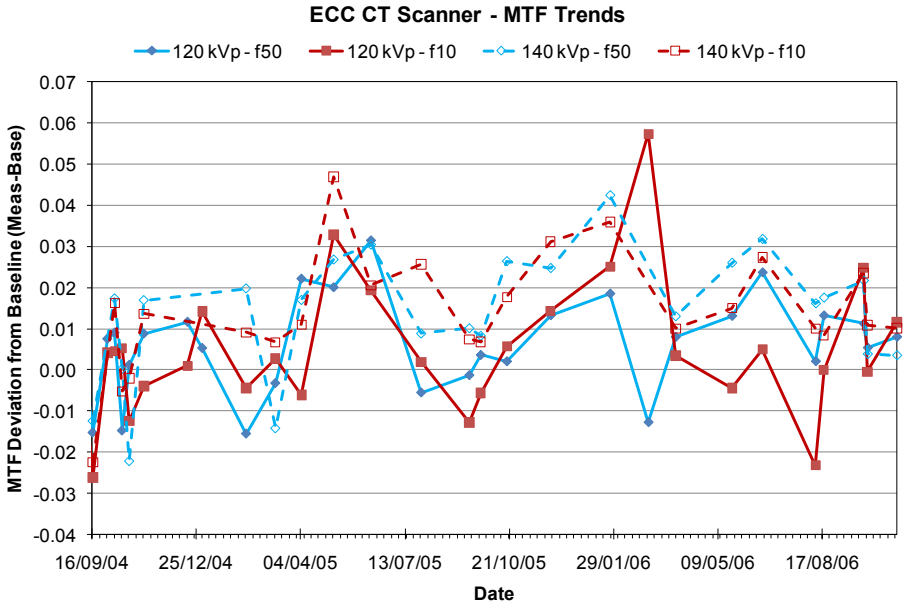


Figure 4.12. – f_{50} and f_{10} modulation transfer function (MTF) time-trends of the radiotherapy CT scanner in Edinburgh Cancer Centre (ECC), measured using the Varian CT Performance Phantom.

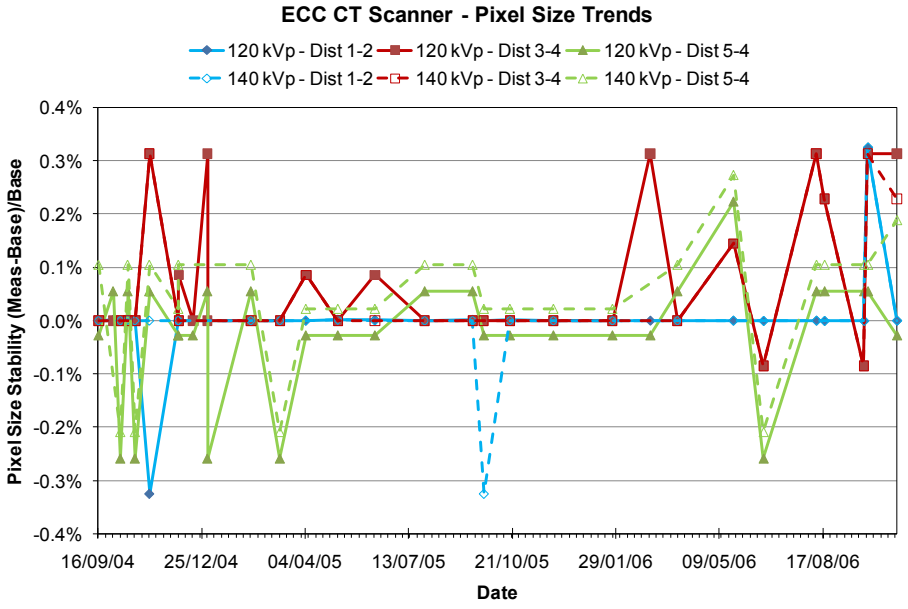


Figure 4.13. – Pixel size time-trends of the radiotherapy CT scanner in Edinburgh Cancer Centre (ECC), measured using the Varian CT Performance Phantom.

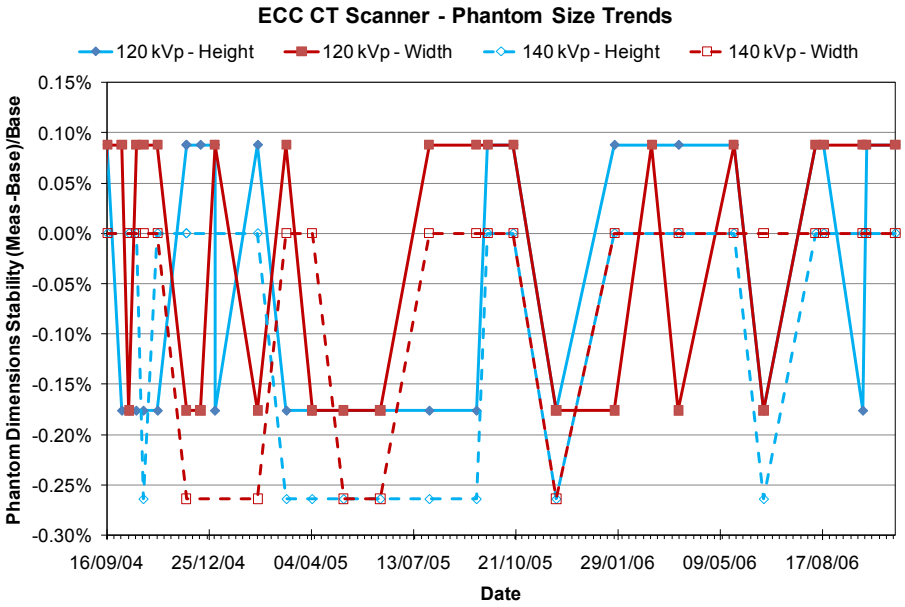


Figure 4.14. – Time-trends of the height and width of the Varian CT Performance phantom scanned on the radiotherapy CT scanner in Edinburgh Cancer Centre (ECC).

4.7. Catphan Series of Phantoms

Phantom Laboratory's 'Catphan' series of phantoms has gained widespread acceptance as the 'gold standard' for CT performance evaluation, often being employed in both diagnostic and radiotherapy imaging for detailed system characterisation or to define reference performance standards (see, for example [151, 253, 345]). Three models of the Catphan are in general use today, the Catphan 500, Catphan 504 and Catphan 600. Common to all models is an outer cylinder into into disc-based analysis modules are inserted, with the various analysis modules offering different test objects to perform a wide range of evaluations. The difference between the phantom models is the type and arrangement of analysis modules included. These are summarised below:

- **Catphan 500**. This is the basic Catphan model designed for use with conventional, single-slice CT scanners. The Catphans 504 and 600 are evolutions of this.

- **Catphan 600.** Optimised for next-generation multi-slice scanners, this replaces the 'Alignment & Sensitometry' module of the Catphan 500 with a similar module that contains 8 instead of 4 electron-density inserts. Furthermore, an additional module is included which contains ramps comprising tiny ball-bearings. These allow a more accurate assessment of slice thickness for multi-slice scanners than the angled ramps included in the 'Alignment & Sensitometry' module.

- **Catphan 504.** A hybrid between the Catphans 500 and 600, this includes the 8 material 'Alignment & Sensitometry' module of the 600 but not the ball-bearing ramps of the 'Multi-Slice Thickness' module. Furthermore, this phantom has been optimised for CBCT applications by rearranging the modules so that the 'Alignment & Sensitometry' module is near the centre of the phantom. This module may used for calibration and verification of a scanner's electron-density conversion curve. Assuming a cone-beam geometry where the beam central-axis is aligned with the centre of the phantom, the scatter conditions at this module will therefore be representative of those in a real patient. That Catphan 504 is included with all Varian and Elekta CBCT systems.

Diagrams illustrating the arrangement of modules in each of the Catphan models are included in figure 4.15 and photographs of the Catphan 504 being utilised in Oxford for CT scanner and OBI CBCT assessment are in figure 4.16.

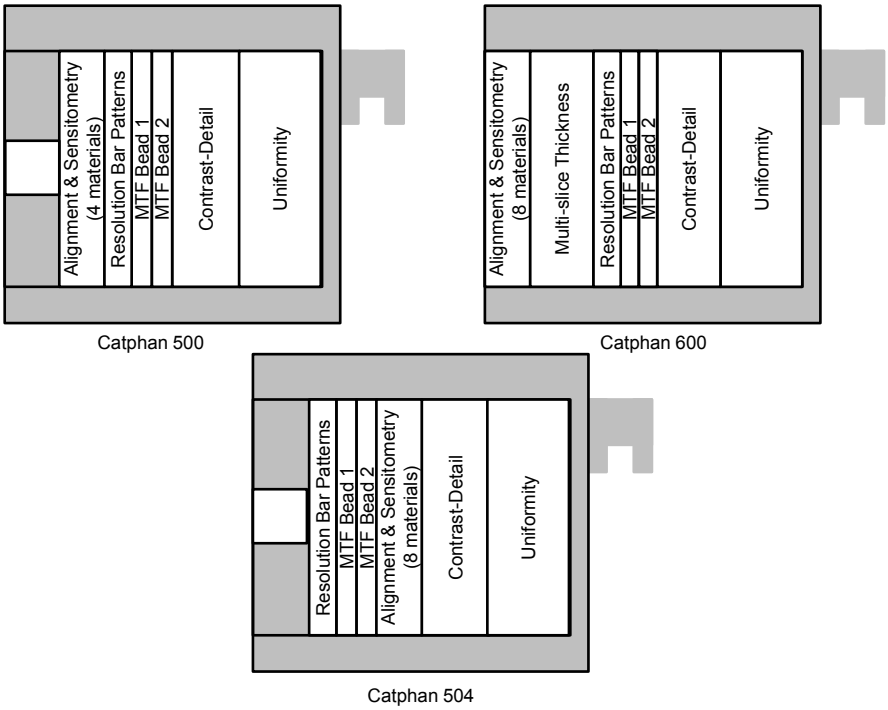


Figure 4.15. – Schematic diagram showing the arrangement of analysis modules in the Catphan 500, 504 and 600 phantoms.

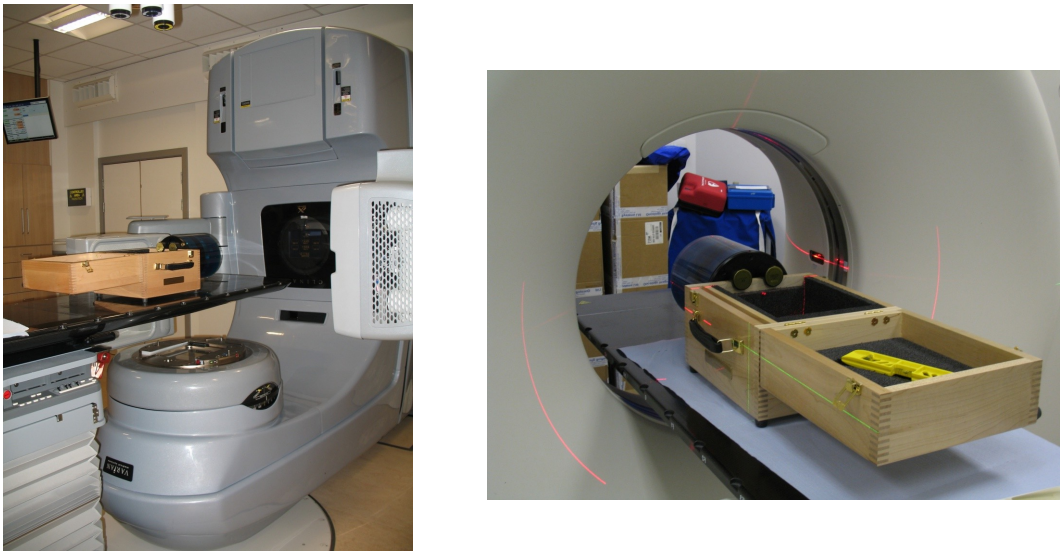


Figure 4.16. – Catphan 504 phantom being scanned on a conventional CT scanner (right) and Varian OBI CBCT unit (left).

4.7.1. Catphan Alignment & Sensitometry Modules

Diagrams and CT scans of a slice through the Catphan 'Alignment & Sensitometry' modules are presented in 4.17, with the modules being identical across all phantoms except for the number, and potentially material, of the sensitometry inserts. Included in the modules are:

- **Sensitometry Inserts 1–4 or 1–8.** These are plugs of materials of calibrated electron-densities, as indicated in the table below. It is noted that the Catphan 504 / 600 module has two voids for measuring air HU. However, because the module is located at the very end of the housing cylinder in the Catphan 600 there is the option with that phantom for a small container to be inserted which can be filled with water through an inlet on the edge of the phantom.

- **Geometric Linearity Holes H1–H4.** Distances can be measured between the holes to verify geometric linearity. In the standard phantom configuration one of the holes is filled with a PTFE rod. By default this is H2 but the end-user is free to move the rod to a new position. The rod can be utilised to verify phantom orientation or to distinguish between different phantoms of the same model.

- **23° Angled Ramps.** For measuring slice thickness and phantom alignment. When the phantom is aligned so that a slice intersects the mid-plane of the module the bright lines from the rod appear centred on the X and Y axes. However, if the slice is offset perpendicular to the module (i.e. in the direction of couch travel) the bright lines all appear to move clockwise or anti-clockwise around a square. Depending upon requirements, the ramps can either be used to identify the actual location of the slice relative to the centre of the module (by calculating the positions of the centres of each ramp and compensating for the ramp angle), or as an alignment aid by zeroing the scanner on the slice where the ramps appear centred.

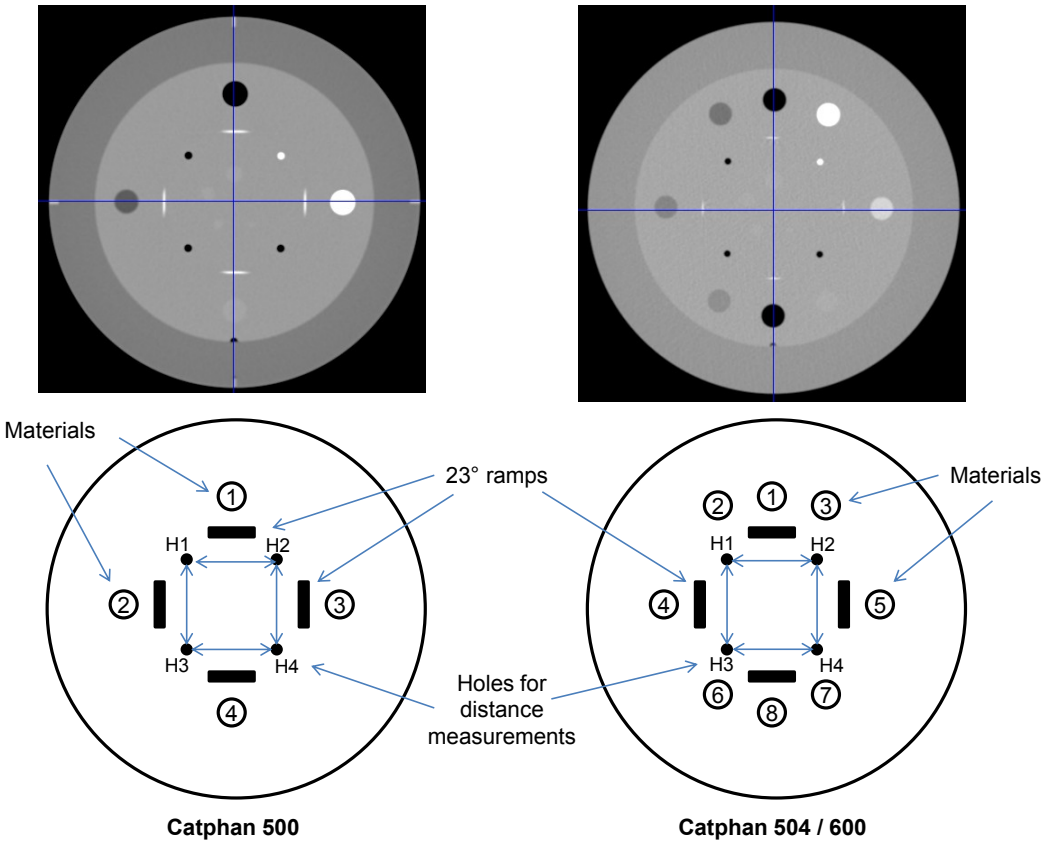
- **Sub-Slice Contrast-Detail Pattern.** Not shown in the schematic diagrams, but visible at the centres of the CT slices, this can be interpreted visually to provide a subjective assessment of scanner performance.

A comprehensive analysis tree was developed in IQWorks. This comprised:

- Identification of the outer edge of the phantom, including calculating of phantom width, height and the centre of extremes of the edge. This was used both to localise other ROIs and evaluate geometric linearity.

- Calculation of the mean and standard deviation in ROIs placed on each electron density insert. These results are plotted as a function of electron density.
- Measurement of the distances between each pair of holes: H1-H2, H2-H4, H3-H4 and H1-H3. If the FOV and hence pixel size are small, so that each hole is spread over a number of pixels, then the holes are localised by edge detection and the COE of the edge taken as the centre of a hole. If this is not possible, then the location of the minimum pixel value (or maximum for the hole containing the PTFE rod) within the search ROI is used. Each distance is compared against the expected the expected value with a tolerance of ±0.5 mm.
- Calculation of contrast (using equation 2.8) and CNR (using equation 2.5) between all sensitometry inserts and the module base material.
- Calculation of slice thickness using each of the angled ramps.

It should be noted that separate trees were required for the Catphan 504 and 600 phantoms, even though the configuration of the module is the same for each phantom model. This was because the whole module appears to be rotated through 180° in the Catphan 504 phantom, although this is not documented scanner in the manual[105].



Material Number			Material	Electron Density Rel to Water	Nominal HU
Catphan 504	Catphan 600	Catphan 500		ρ_e'	
1	1	1	Air	0.000	-1000
2	2		PMP	0.853	-200
3	3	3	PTFE	1.867	990
4	4	2	LDPE	0.944	-100
5	5		POM	1.353	340
6	6		Polystyrene	1.017	-35
7	7	4	Acrylic	1.146	120
8	8		Air	0.000	-1000
	8		Water	1.000	0

Figure 4.17. – CT scans (top) and schematic diagrams (middle) of the Catphan 500, 504 and 600 alignment modules. Bottom–Table indicating the arrangement and electron densities of the sensitometry inserts in each phantom. In the table the acronyms PTFE, LDPE and POM stand for PolyTetraFluoroEthylene, Low Density PolyEthylene and PolyOxyMethylene respectively.

Scans of a Catphan 500 in Edinburgh and a Catphan 504 in Oxford were acquired using protocols E12–E14 (in Edinburgh) and O3–O5 and O8 (in Oxford), described in tables 4.2 and 4.3. Although the scanning protocols could not be identical between the two centres because the scans were performed at different points in time, the results of the previous sections suggest that only the kVp should have a significant impact on sensitometry measurements. These protocols all employed scan and reconstruction fields of view of 25 cm, optimised for the 20 cm diameter of the Catphan. In Oxford additional scans were acquired at SFOV / DFOV = 50 cm / 50 cm (protocol O6) and SFOV / DFOV = 50 cm / 65 cm (protocol O7) to specifically evaluate the larger FOV modes.

Measured electron-density conversion curves are compared against the ICRU-42 and Catphan nominal curves in figure 4.18. In the error plot the results are plotted relative to the Catphan nominal values. As before, the tolerance band calculated to maintain dose calculation accuracy of ±2% is present, but this time the ±40 HU tolerance band stated by Varian in the specification of their OBI and Acuity equipment[344–346] is also included.

Somewhat surprisingly, the trends exhibited previously are not observed in these assessments. Furthermore, the influence of kVp is far less pronounced, with all measurements at all kVps falling within both tolerance bands across the entire electron-density range. After some consideration, it was concluded that this was due to the electron-density inserts in the Catphan module not being tissue-equivalent. Specifically, whereas radiologically tissue-equivalent cortical bone has sharp absorption edges in the low keV energy range, the linear attenuation coefficient of PTFE varies smoothly over the same range. This is illustrated in figure 4.19, which also shows that a less-dense Catphan material, polyethylene, is a closer match to the similarly less dense tissue-equivalent ICRU 44 adipose tissue[132, 143].

Although this finding is the reason why the Catphan nominal curve is different from that of ICRU 42 curve, these are important and interesting results which are not dealt with in the radiotherapy imaging literature. Firstly, whilst the Catphan can certainly be utilised to calibrate the HU scale of a given CT system, it cannot be used to verify the HU electron-density conversion curve for biologically representative materials. These results demonstrate that it is insufficient simply to check that the Catphan sensitometry curve is as expected. Furthermore, if the Catphan is to be relied upon for regular constancy checks, as suggested in manufacturers' guidance[345] and in the scientific literature[359],

it is crucial that changes in performance which would affect the pixel values of biologically representative materials are also also reflected in those of the materials provided with the Catphan. However, given the very different trends found above it is not at all certain that this would be the case and further investigation is required.

It is interesting that CT phantoms commonly utilised in the diagnostic sphere incorporate engineering materials such as PMMA, polyethylene and PTFE for assessing imaging performance (see, for example [26]), sometimes alongside tissue-equivalent materials (e.g. in [2]). In particular, as in the Catphan, PTFE is frequently found as a substitute for bone, presumably because in CT scans it returns a HU roughly similar to that of dense bone (of the order of 1000 HU), provides a high contrast object for optimisation and QA experiments and is a material which is relatively straightforward and inexpensive to manufacture and machine. Use of such materials may be appropriate in diagnostic imaging because the clinical task involves identifying structures of different contrast levels against a noisy background, so is concerned more with detectability and reproducibility of signal level (to maintain a constant system transfer curve) rather than the determination of absolute material properties. However, even so, the results found above suggest that a system transfer curve calculated using materials with linear attenuation coefficients that vary smoothly with energy over the 10-100 keV range may not be sensitive to changes in beam spectrum. Therefore, it is proposed that – even for diagnostic imaging applications – inclusion of some tissue equivalent materials in CT phantoms may be wise.

In a radiotherapy context, it is strongly recommended that the Catphan 'Alignment & Sensitometry' module not be relied upon as the primary phantom for verifying the HU to electron-density calibration curve.

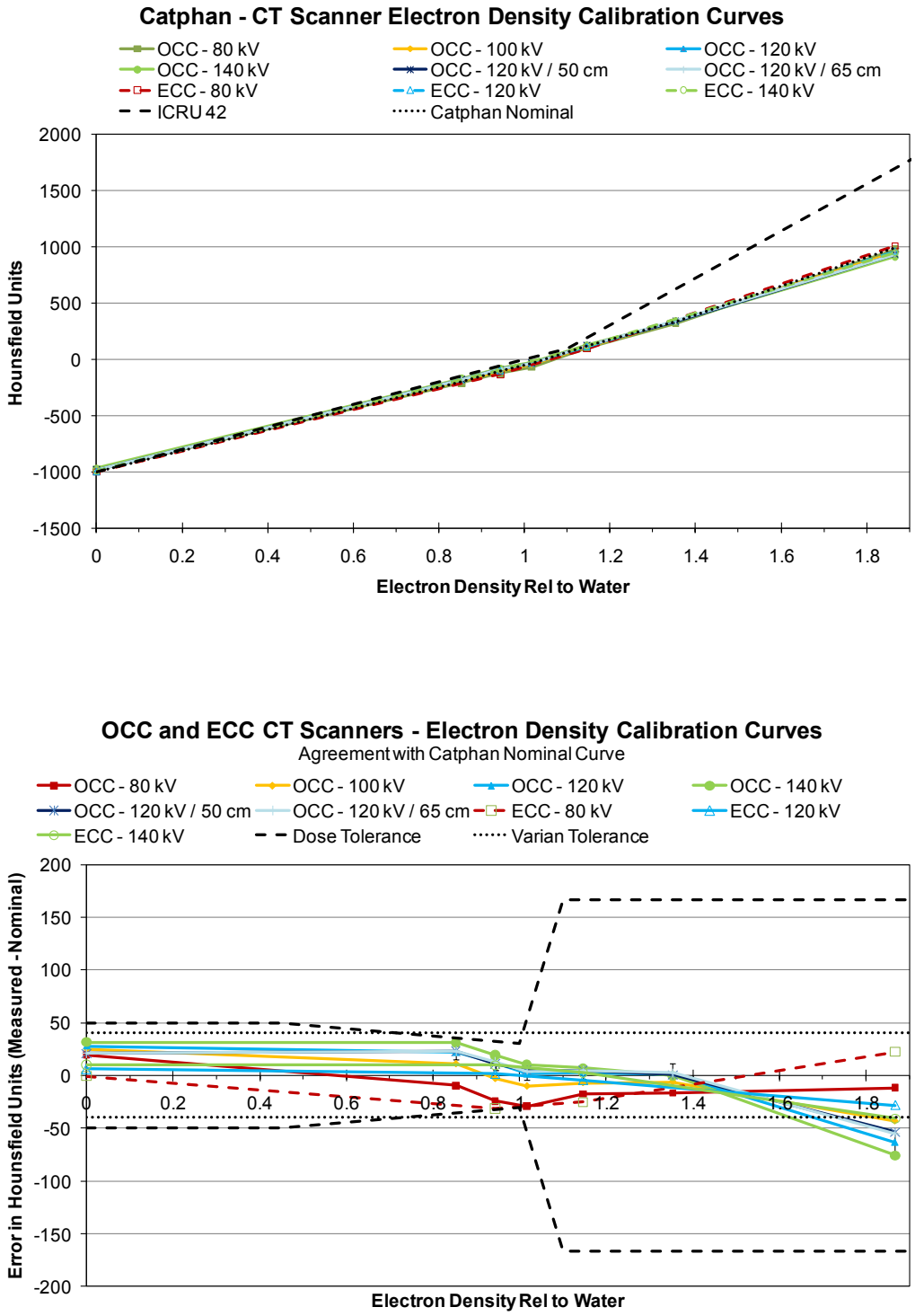


Figure 4.18. – Electron density calibration curves of radiotherapy CT scanners in Oxford and Edinburgh, measured using the Catphan 500 / 504 phantoms and IQWorks. Error bars are ±1 SD of the pixel values in each ROI. Top–Measured calibration curves, with Catphan nominal curve and ICRU-42 curve present as references. Bottom–Deviation of measured curves from Catphan nominal.

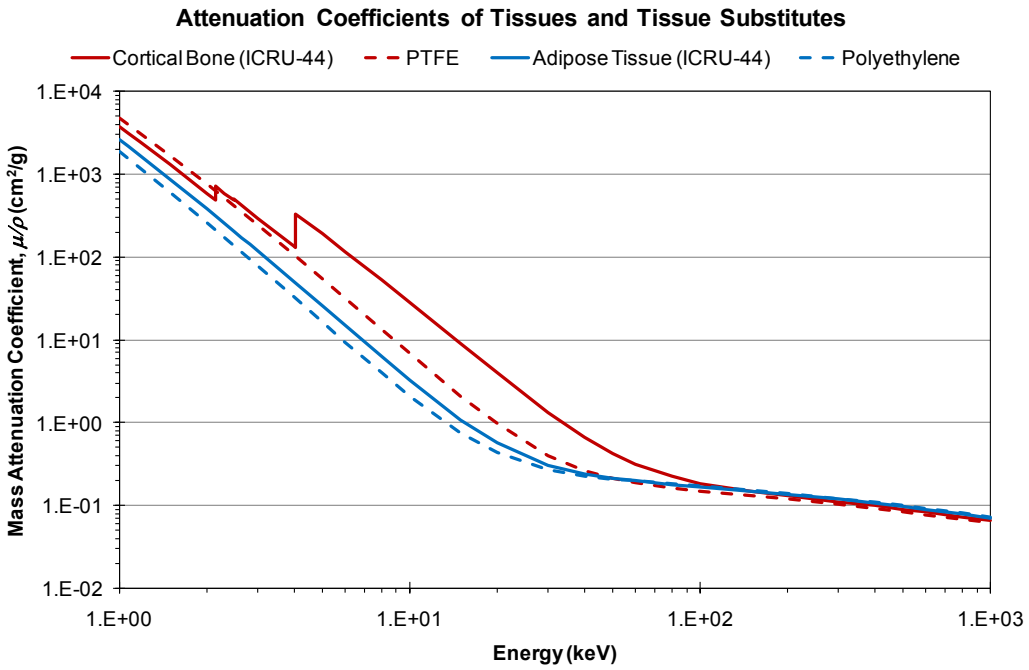


Figure 4.19. – Plot of mass attenuation coefficient μ/ρ over the diagnostic imaging energy range for ICRU-44 defined tissues (cortical bone and adipose tissue)[143] and Catphan tissue substitutes (PTFE and polyethylene)[105]. Data are those of Hubbell and Seltzer[132].

In Edinburgh the Catphan 500 was scanned regularly as part of the CT scanner QA programme, during the same sessions as the Varian Performance Phantom. Protocols E12–E14 in table 4.2 were utilised for these measurements.

A plot of sensitometry over time is shown in 4.20 for four materials: LDPE, acrylic, PTFE and the phantom base material. Although the trends are similar to those observed with the Varian Performance Phantom, namely that all materials generally agree with baseline, at both 120 kVp and 140 kVp, to within ±5 HU, there are clear spikes in the results. These are due to the X-ray tube arcing (or "spitting") due to a gradual breakdown of the vacuum in the tube as it neared the end of its life. Indeed, there are 140 kVp data points missing from the plot because there were occasions when the 140 kVp mode would fail to operate. It is interesting that the same spikes were not seen in the comparable measurements made over the same time frame using the Varian Performance Phantom. It is thought this is due to a scan of the longer volume of the Catphan being a greater load on the X-ray tube, as well as the $CTDI_{vol}$ for individual slices being considerably higher (more than double, in comparison with the Varian phantom protocols E10 and E11 in table 4.2). This demonstrates the importance of the choice of scanning protocol for QA purposes. Ideally, alongside typical acquisition settings a scan length representative of that used with patients should be imaged on a regular basis, although in practice this may not be possible.

Figure 4.21 illustrates that CNR results were also influenced by tube spits, with spits resulting in a significant spike in image noise and a corresponding dip in CNR. Otherwise, the CNR results lay within a band ±20% about baseline. Direct comparisons cannot be drawn between the absolute CNR measurements performed using the two phantoms because of the different attenuation properties of the materials in each phantom. However, it is surprising that the CNR results fluctuate over a larger range (±20%) than observed for the Varian Performance Phantom (±10%), especially given that the $CTDI_{vol}$s for the Catphan measurements (55.6 mGy at 120 kVp and 74.7 mGy at 140 kVp) were approximately 3.5 times those of the Varian phantom (16.6 mGy and 22.9 mGy, respectively), because the noise present should be less. A possible explanation might be that because the noise is less any fluctuation in what is a small numerical value on the denominator of the CNR calculation would result in a more significant change in CNR, although further work is required to investigate this.

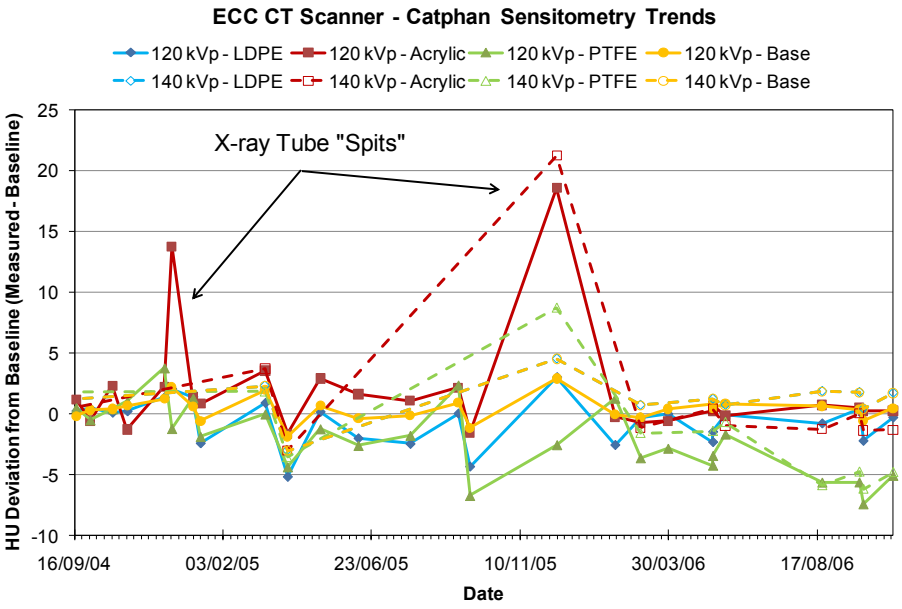


Figure 4.20. – Sensitometry time-trends of the radiotherapy CT scanner in Edinburgh Cancer Centre (ECC). Data are presented for the sensitometry materials in the Catphan 500 phantom scanned at 120 and 140 kVp.

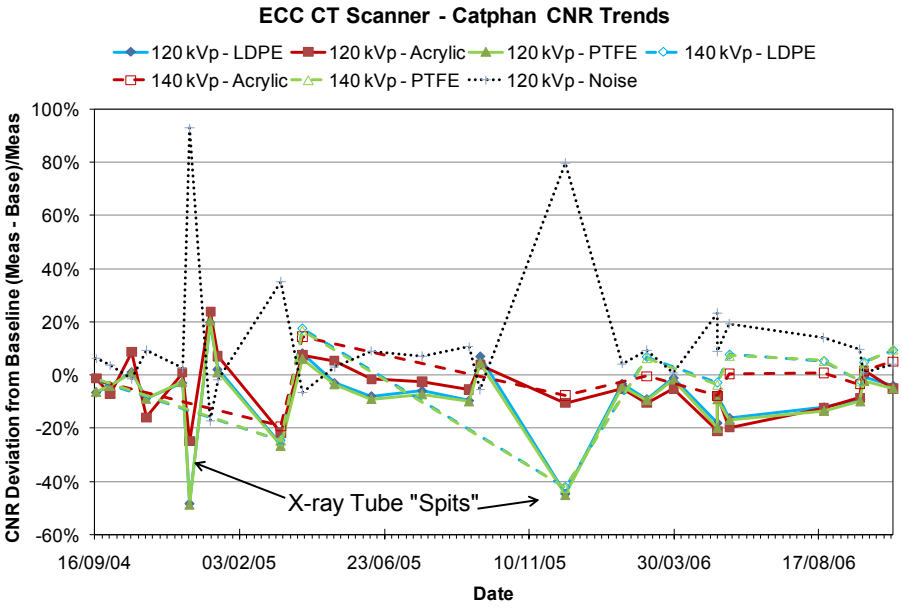


Figure 4.21. – Contrast-to-noise ratio (CNR) time-trends of the radiotherapy CT scanner in Edinburgh Cancer Centre (ECC). Data are presented for the sensitometry materials in the Catphan 500 phantom scanned at 120 and 140 kVp.

4.7.2. Catphan MTF Bead Modules

Both 'MTF Bead' modules comprise a high-density tungsten carbide ball of diameter 0.28 mm embedded in a uniform material[105]. The modules are identical except that in one the bead is in the upper quadrant of the image slice and in the other it is in the lower.

An experiment was performed using the Oxford scanner to investigate the influence of field of view on spatial resolution. Three scans were performed at 120 kVp using the standard abdomen protocol but with acquisition fields of view of 25 cm, 50 cm and 65 cm (protocols O5–O7 in table 4.3). From the calculated MTF curves in figure 4.22 is evident that MTF deteriorates significantly with increasing FOV. This result is as expected due to the increase in pixel size with FOV.

During an ImPACT evaluation of this scanner[175] an f_{50} of 0.38 cycles / mm was measured for a similar scan protocol using SFOV / DFOV = 50 cm / 38 cm. This is in good agreement with the current work because it lies between the results measured here at 25 cm (f_{50} = 0.40 cycles /mm) and 50 cm (f_{50} = 0.33 cycles /mm), as should be expected. Furthermore, the proportionate degradation in f_{50} between 50 cm and 65 cm is of the order of 80%, which is consistent with an increase in pixel size of 77%.

Although these results are as expected it is still useful to be able to quantify the effect. There is a tendency in radiotherapy CT scanning to always use the same standard protocol for a given treatment site without fine-tuning it to the patient imaging scenario. For example, if the scanner is set to acquire a 65 cm FOV scan for a breast patient then the diameter is generally not reduced to 50 cm for thinner patients or those not requiring highly-angled breast boards. However, these results provide strong evidence that a potentially relevant improvement in spatial resolution can be gained by consciously optimising the FOV to patient size.

It is also worth noting that the HiSpeed QX/i MTF curve for the 50 cm FOV experiment is superior to those in figure 4.11 measured under similar conditions for the HiSpeed FX/i scanner. This is consistent with the older FX/i model having larger detector spacing and slower sampling electronics, as specified in table 4.1, and is also in agreement with the measurements of other researchers[45, 175].

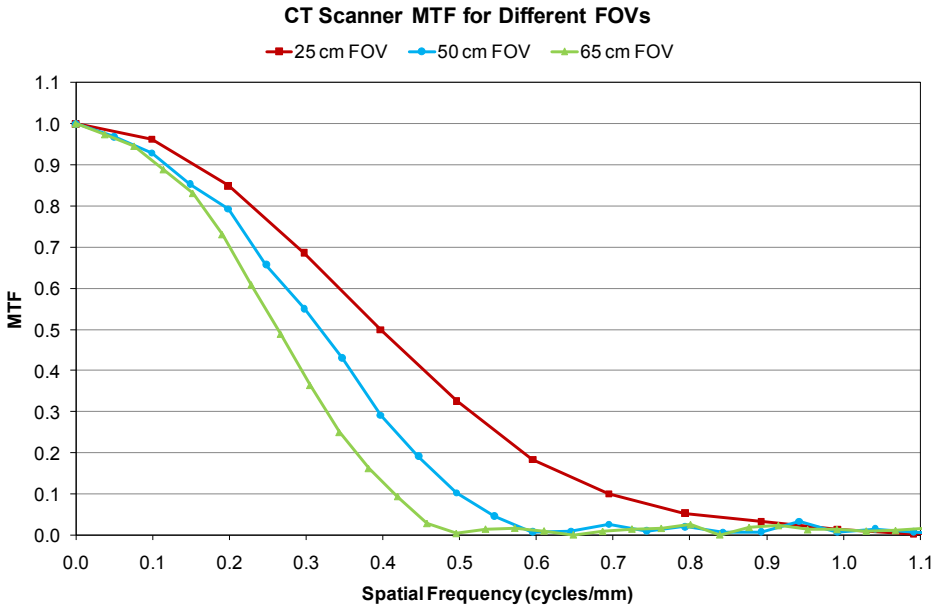


Figure 4.22. – Modulation transfer function (MTF) of the radiotherapy CT scanner in Oxford Cancer Centre (OCC) operating at three different fields of view, measured using the Catphan impulse bead and IQWorks.

4.7.3. Catphan Uniformity Module

Being a simple cylinder of uniform material, the Catphan 'Uniformity' module can be used to calculate a wide range of uniformity metrics. The IQWorks analysis tree developed for this module is illustrated in figure 4.23, and contains:

- Application of a Signal Calibrator to effectively linearise the CT data by adding 1000 to all pixel values. This is necessary for the calculation of absolute uniformity metrics.
- Phantom edge detection to aid placement of other ROIs.
- Circular ROIs of diameter 13 cm and 18 cm, defining 'inner' and 'outer' FOVs respectively. The inner FOV is constrained to the Uniformity module itself whilst the outer FOV also encompasses the cylindrical housing of the Catphan. Depending upon the scanner and imaging scenario, including the housing in a uniformity calculation may be acceptable because its pixel values lie within 10-15 HU of those of the Uniformity module, well within the ±40 HU uniformity tolerance specified by Varian for the Acuity and OBI CBCT systems[345]. Furthermore, if comparative results are reported relative to a baseline measurement then any non-uniformity caused by a step change at the interface between the module and housing becomes irrelevant. Basic statistics and the coefficient of variation are calculated for each ROI.
- Horizontal, vertical and diagonal line profiles through the centre of the phantom, covering both the inner and outer FOVs. For the horizontal and vertical profiles these are calculated as the average of 7 adjacent rows or columns. Only a single lines are included for the diagonal profiles. A range of uniformity metrics is calculated for each profile: integral, integral+, integral- and differential uniformity, and the coefficient of variation.
- 4 square ROIs of side 1 cm are placed in the 'north', 'south', 'east' and 'west' positions at the edges of the inner and outer FOVs. A further ROI of the same dimensions is located at the centre of the phantom. These are used to calculate the 'maximum difference' and 'U1' metrics specified in section B.15.

With so many metrics being calculated it is easy to cause confusion by trying to interpret too many results simultaneously. It is therefore important to be able to reduce these into a manageable subset which is still sensitive enough

to demonstrate changes in performance. If issues arise or trends are identified in this subset then the full battery of metrics can be consulted to investigate further. It was found that a useful set of metrics to monitor regularly included:

- Integral and differential uniformity for all line profiles, with the worst case results being reported as a single metric for each.
- Coefficients of variation for the circular ROIs.
- U1 for each FOV.

In summary, it is suggested that in the first instance only 4 uniformity metrics are considered for each of the inner and outer FOVs. Indeed, it is arguable that although the integral and differential uniformity measures are reduced metrics they are actually more sensitive overall because they represent worst-case results.

Outer FOV uniformity metrics plotted over time for the 120 kVp Edinburgh QA scans described earlier (protocol E13 in table 4.2) are presented in figure 4.24, in which the 'Global Uniformity' metric corresponds to U1. Except for two clear spikes, all metrics are within ±0.7% of baseline across the whole time period. Following an investigation it transpired that the spikes were due to a shadow appearing at the base of the image resulting from the phantom being positioned over a part of the couch where there was an underlying support structure. It is reassuring that the differential, integral and coefficient of variation metrics were sufficiently sensitive to identify this artefact. However, the global (U1) metric did not respond at all to the presence of the artefact and this is of concern because it is this metric which tends to form the basis of performance recommendations[156].

From the results of this work it is recommended that U1 is not relied upon as the only measure of image uniformity, but that instead the four metrics described above are routinely considered.

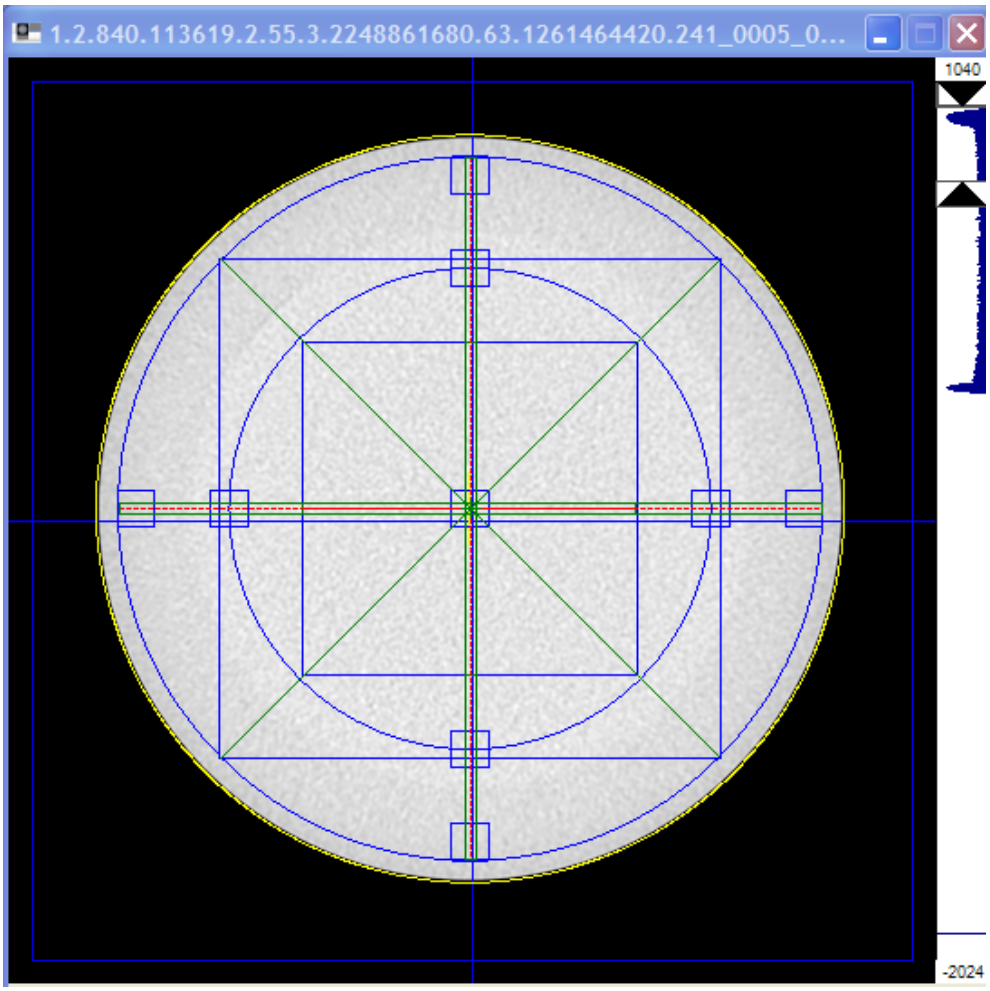


Figure 4.23. – IQWorks screenshot illustrating analysis of the Catphan uniformity module. This demonstrates the inner and outer circular regions of interest (ROIs), 7 localised square ROIs (1 central, 4 at the top, bottom, left and right of the inner circle, and 4 around the peripheral circle) and 4 line profiles (horizontal, vertical and the two diagonals).

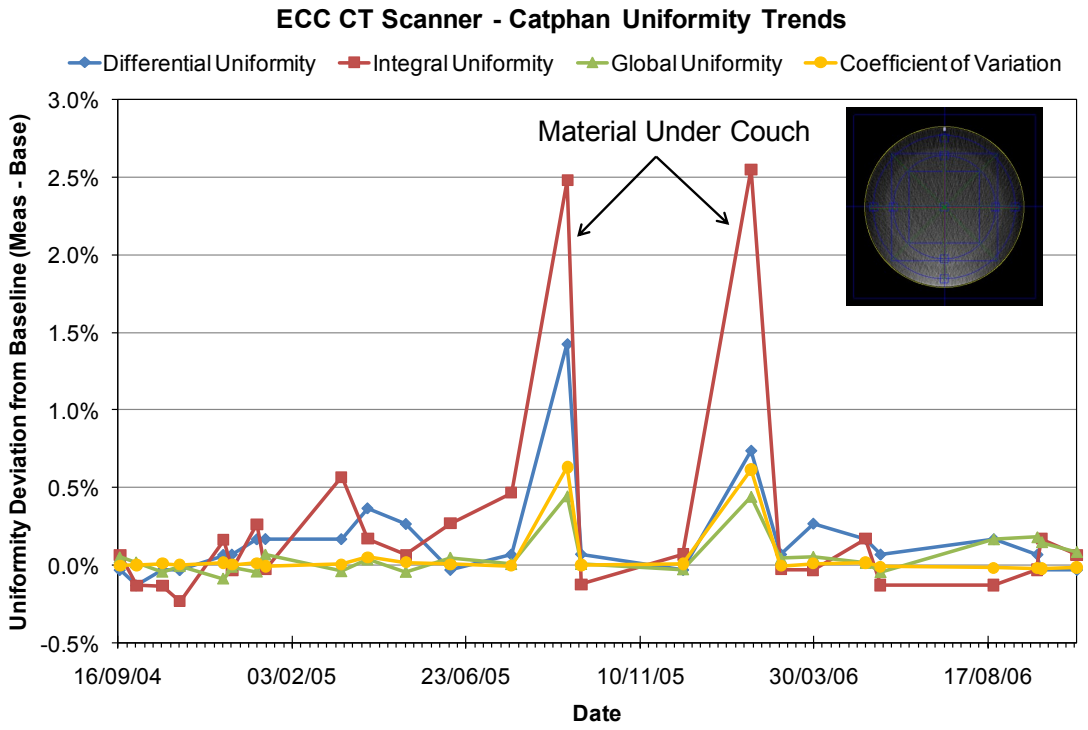


Figure 4.24. – Uniformity time-trends of the radiotherapy CT scanner in Edinburgh Cancer Centre (ECC). Values of four different uniformity indices are presented relative to the commissioning baseline. The insert illustrates a gradient-like artefact caused by structural materials present under the longitudinal extremity of the couch-top.

4.8. Practical Application: Commissioning of CBCT Systems

4.8.1. Overview

This section discusses the application of the IQWorks performance evaluation framework to the commissioning of two Varian OBI CBCT units in the new Oxford Cancer Centre. A rigorous alignment process coupled with regular QA and geometrical calibration ensure that the isocentres of the kilovoltage and megavoltage systems coincide to within a sphere of radius 1.5 mm[344, 359], thus enabling images representative of the treatment geometry to be acquired using the kilovoltage system. Essentially, during installation the isocentres of the two beamlines are accurately characterised through a combination of optical and radiographic checks. If there is a discrepancy, then the kilovoltage beamline is shimmed before the robotic arms are permanently mounted to the linac base frame. Active feedback from the X-ray tube and kilovoltage detector motion controllers then ensures that any deviations in position during a CBCT scan are automatically corrected, thus helping to maintain the tight isocentre. Furthermore, any residual sag is taken into account through application of a sag correction map which is determined by analysing projection images of a geometrical phantom at different gantry angles[226, 345]. Both the correction map and robotic arm control systems can be adjusted if the isocentres drift out of alignment over time.

Central to the OBI system is the performance of the flat panel detector. This is a PaxScan 4030CB unit which is a 40 cm $\times$ 30 cm amorphous silicon TFT array with a caesium iodide phosphor and was described in detail in section 4.2.

4.8.2. Electron Density Calibration Curve – Catphan 504

Images of the Catphan 504 were acquired on both Oxford OBI units using all the available standard protocols, choosing a slice thickness of 2.5 mm and image matrix of 512 $\times$ 512. Technique settings for these protocols are listed in table 4.4, along with the abbreviations which will be used to distinguish them in the following sections. IQWorks was employed to generate electron density conversion curves 'Alignment & Sensitometry'' module and these are compared against the Catphan nominal curve in figure 4.25. As stated earlier, Varian's specified tolerance for the mean pixel value in ROIs placed on sensitometry inserts is ±40 HU about the nominal[345] and this band is included in the figure

for the purposes of comparison.

It was immediately clear from the results that all but two of the scan protocols (SD Head and Pelvis) were outside specification, some considerably so. This was truly surprising because manufacturer configuration had been completed and both OBI units had formally passed the manufacturer's acceptance test. However, it appears that only one full-fan protocol (SD Head) and one half-fan protocol (Pelvis) are calibrated by the manufacturer during installation and indeed the acceptance test procedure involves only analysing images for a single mode (SD Head). This was a very important finding for a number of reasons. Firstly, the fact that not all modes are calibrated initially is poorly documented. In addition, from practical experience one cannot assume that if any particular mode is within specification it must follow that the others are too. This has implications for the level of testing performed on a regular basis as part of a routine QA programme. Furthermore, although the HU response curve tends to be considered in radiotherapy from the perspective of its implications on dose calculations, the curve also influences the relative contrast between objects of different densities. In figure 4.25 the uncalibrated curves are non-linear and all have slightly different trends. Therefore, patient images acquired using these would not only have incorrect absolute pixel values (potentially of the order of hundreds of HU out) but the contrast observed between similar structures would also be different from one scan protocol to the next.

Following a full recalibration all scan protocols were found to be well within specification, as shown in figure 4.26. For the purposes of comparison, the curves measured for all clinical imaging modes on the Acuity system in Oxford, determined using the same phantom and IQWorks analysis tree, are also included in figure 4.27. The technique settings for these modes are included in table 4.5. Although these curves also represent the state of the equipment immediately following installation they are all within specification, with no additional calibrations having been required.

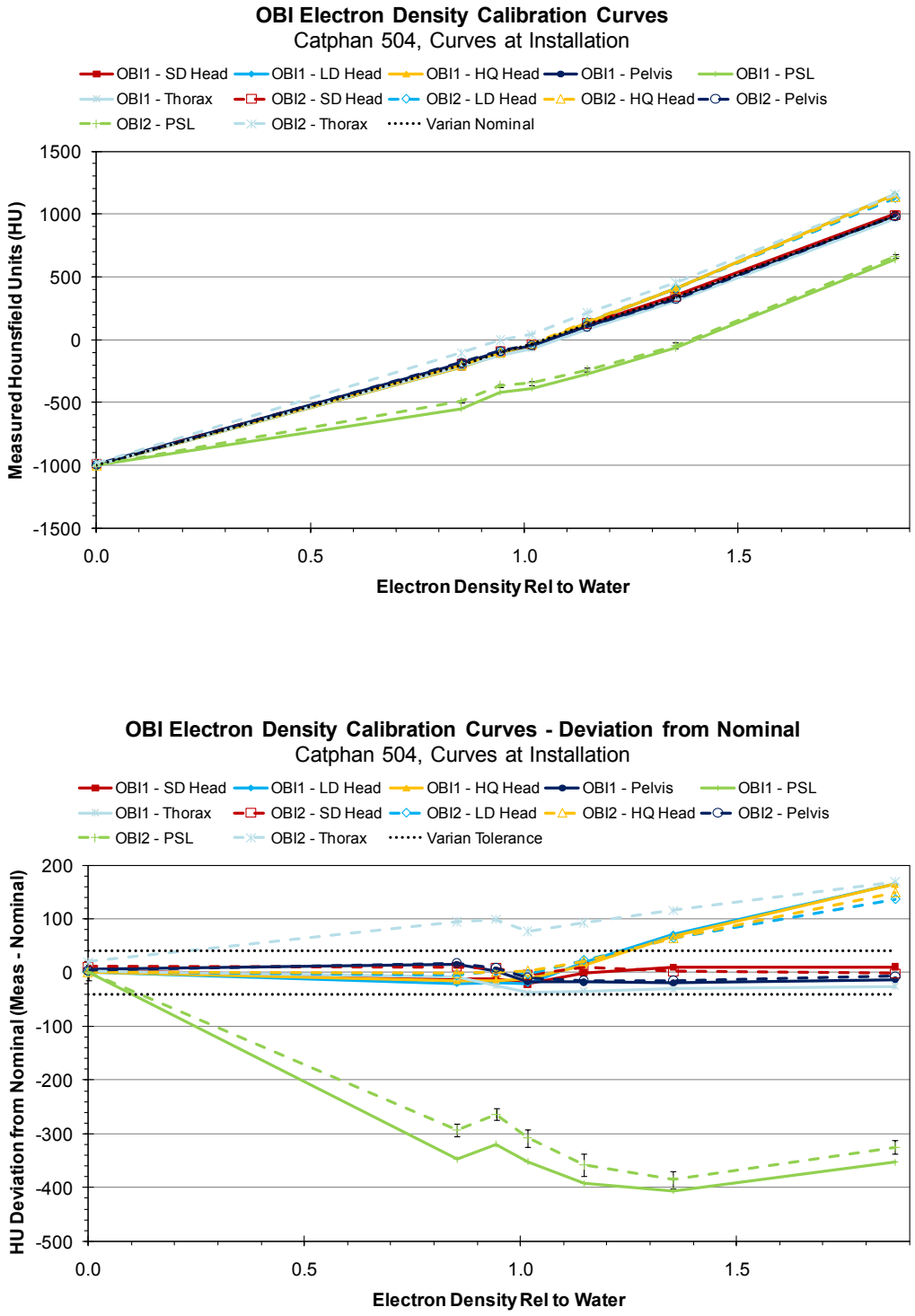


Figure 4.25. – Electron density calibration curves of all clinical imaging modes of two Varian OBI units in Oxford, immediately following installation and acceptance testing. Error bars are ±1 SD of the pixel values in each ROI. Top–Measured calibration curves, with Varian nominal curve present as the ideal reference. Bottom–Deviation of measured curves from Varian nominal. Abbreviations are as in table 4.4.

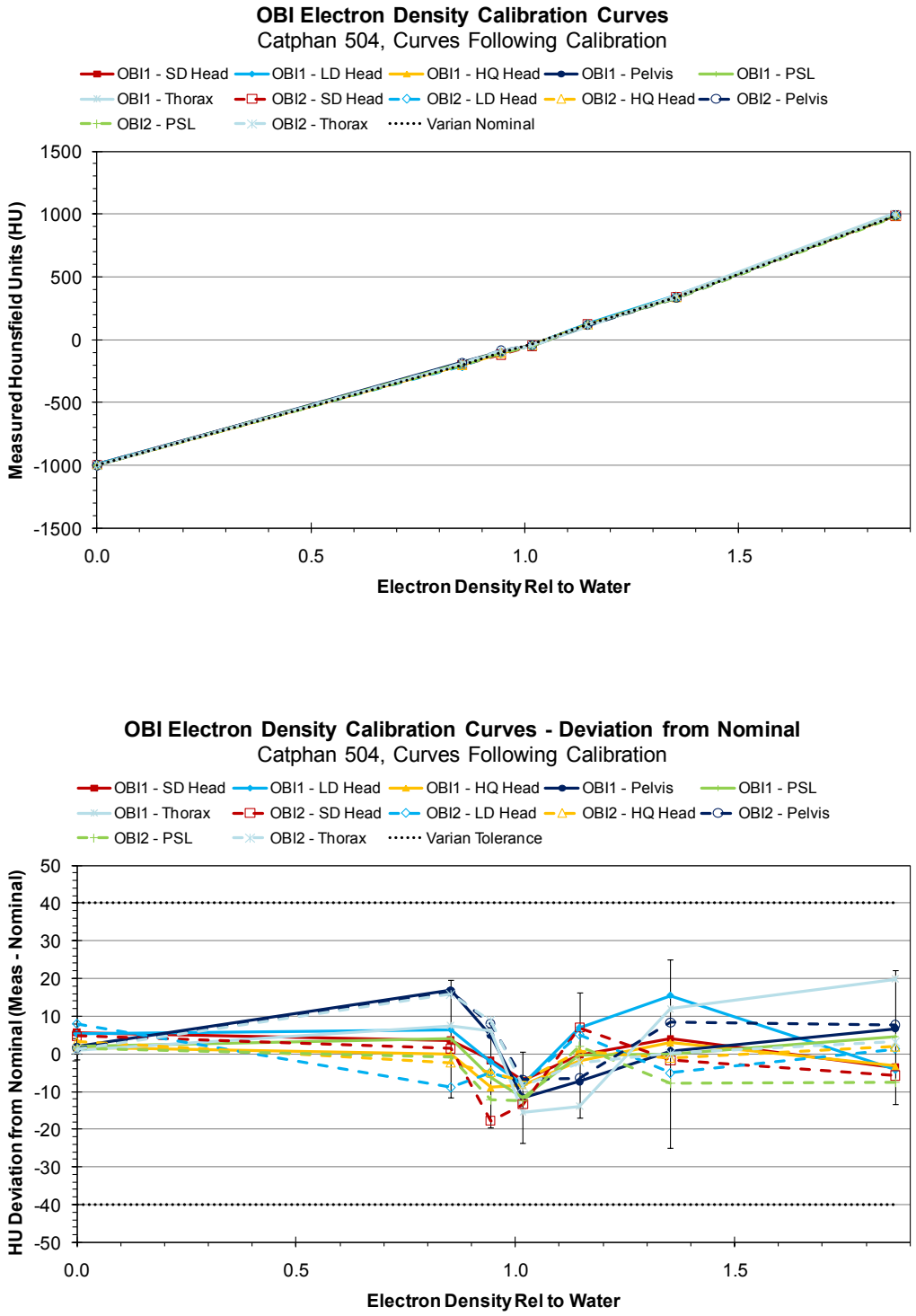


Figure 4.26. – Electron density calibration curves of all clinical imaging modes of two Varian OBI units in Oxford, immediately following comprehensive recalibration. Error bars are ±1 SD of the pixel values in each ROI. Top–Measured calibration curves, with Varian nominal curve present as the ideal reference. Bottom–Deviation of measured curves from Varian nominal. Abbreviations are as in table 4.4.

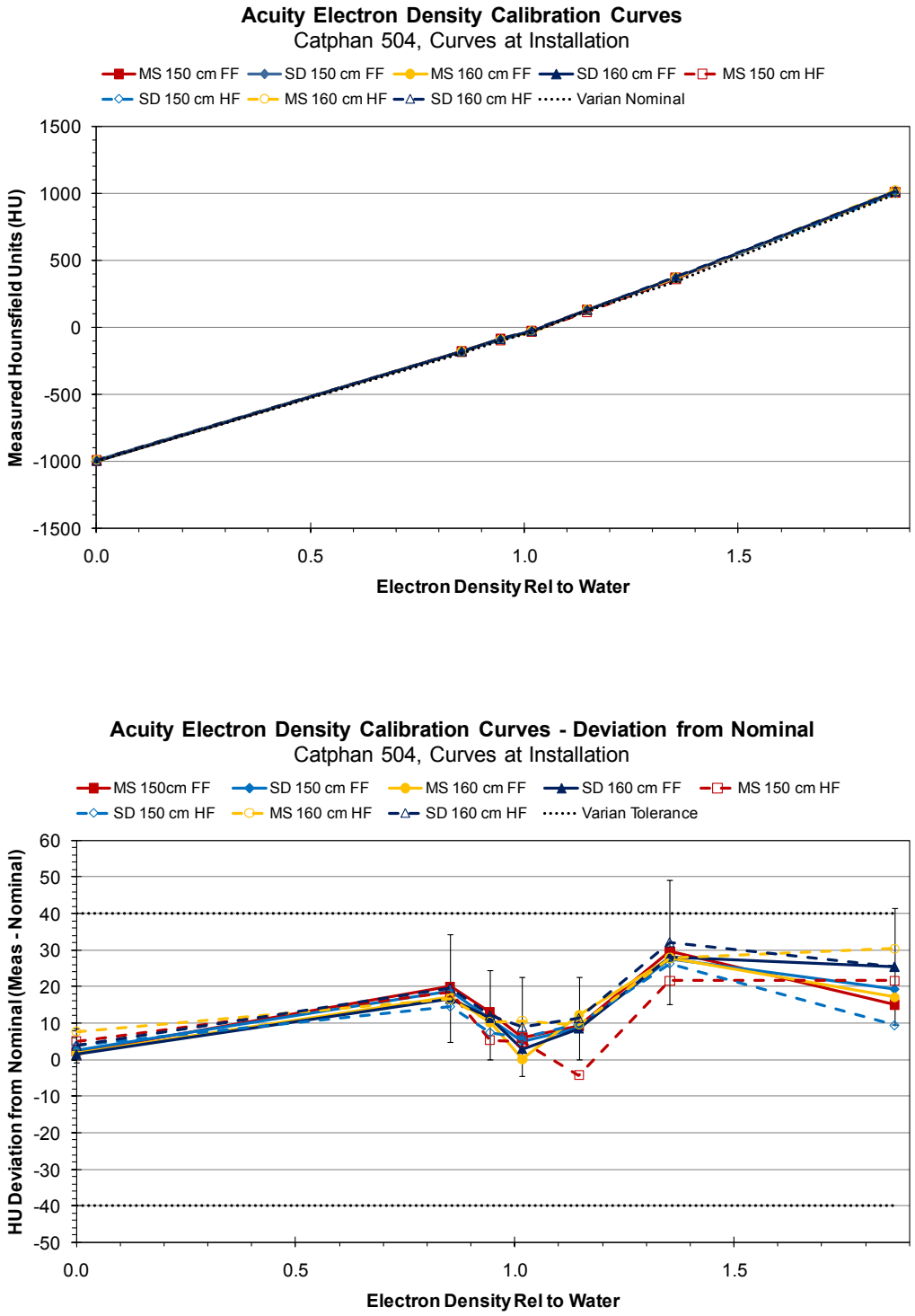


Figure 4.27. – Electron density calibration curves of all clinical imaging modes of the Varian Acuity simulator in Oxford, immediately following installation and acceptance testing. Error bars are ±1 SD of the pixel values in each ROI. Top–Measured calibration curves, with Varian nominal curve present as the ideal reference. Bottom–Deviation of measured curves from Varian nominal. SD and MS refer to Standard Dose and MultiScan respectively.

4.8.3. Electron Density Conversion Curve – Biologically Representative Materials

The importance of using biologically representative tissue-equivalent materials to verify the electron density conversion curve was discussed in section 4.7.1. However, none of the available dedicated electron-density phantoms was suitable for this application because they were all too thin to induce scatter conditions representative of those during calibration conditions. Stacking blocks of solid water either side of the RMI-467 phantom was attempted but phantom alignment became too awkward and the sharp edges of the blocks themselves introduced artefacts into the images.

An alternative approach was to mount a selection of electron-density inserts from the RMI-467 phantom inside a water-filled phantom Nuclear Associates PET-CT phantom (Fluke Biomedical, Ohio, USA). Photographs of the inserts being arranged in the phantom are shown in figure 4.28 and the electron-density curves determined using IQWorks are in 4.29. It is evident from the plotted curves that agreement with the ICRU 42 reference is significantly poorer than would have been expected given the excellent results for the Catphan and the known achievable performance of conventional CT systems measured earlier in this chapter. Indeed, for all modes there are points – and for some modes up to 5 points – which fall outside the tolerance band designated for a ±2% influence on dose calculations.

These results emphasise that for radiotherapy applications involving dose calculations, including basic treatment planning, on-treatment review of dose delivery and advanced adaptive radiotherapy, it is crucial that electron-density conversion curves are verified using appropriate materials. Considering these results in the context of the findings for conventional CT systems it is suggested that it may not be possible to calibrate CT systems to be within specification both for the Catphan and biologically representative materials, and that perhaps biologically representative materials should themselves be used during the calibration process. This is currently being investigated in Oxford and has been followed up with the manufacturer, who appeared to be unaware this was a potential issue.

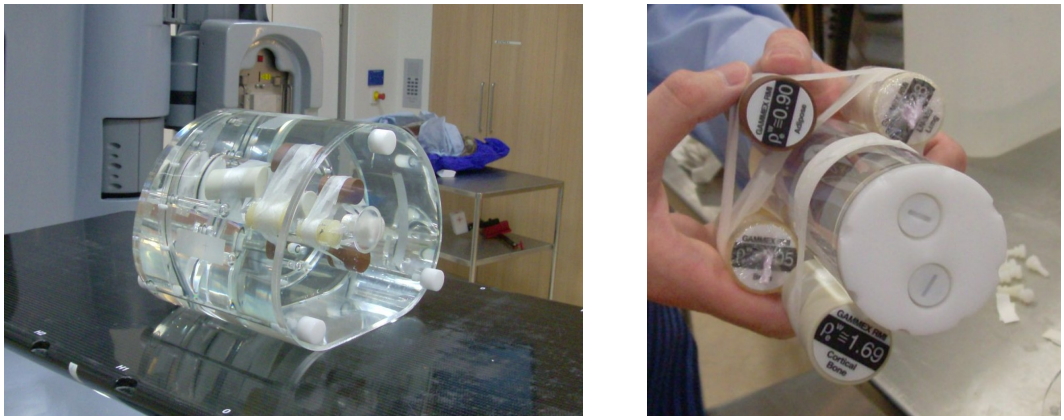


Figure 4.28. – Photograph of the Nuclear Associates PET-CT phantom containing tissue equivalent inserts from the RMI-467 phantom. Left–Phantom filled with pure water and aligned on OBI couch for scanning. Right–Inserts affixed to central column prior to mounting inside phantom.

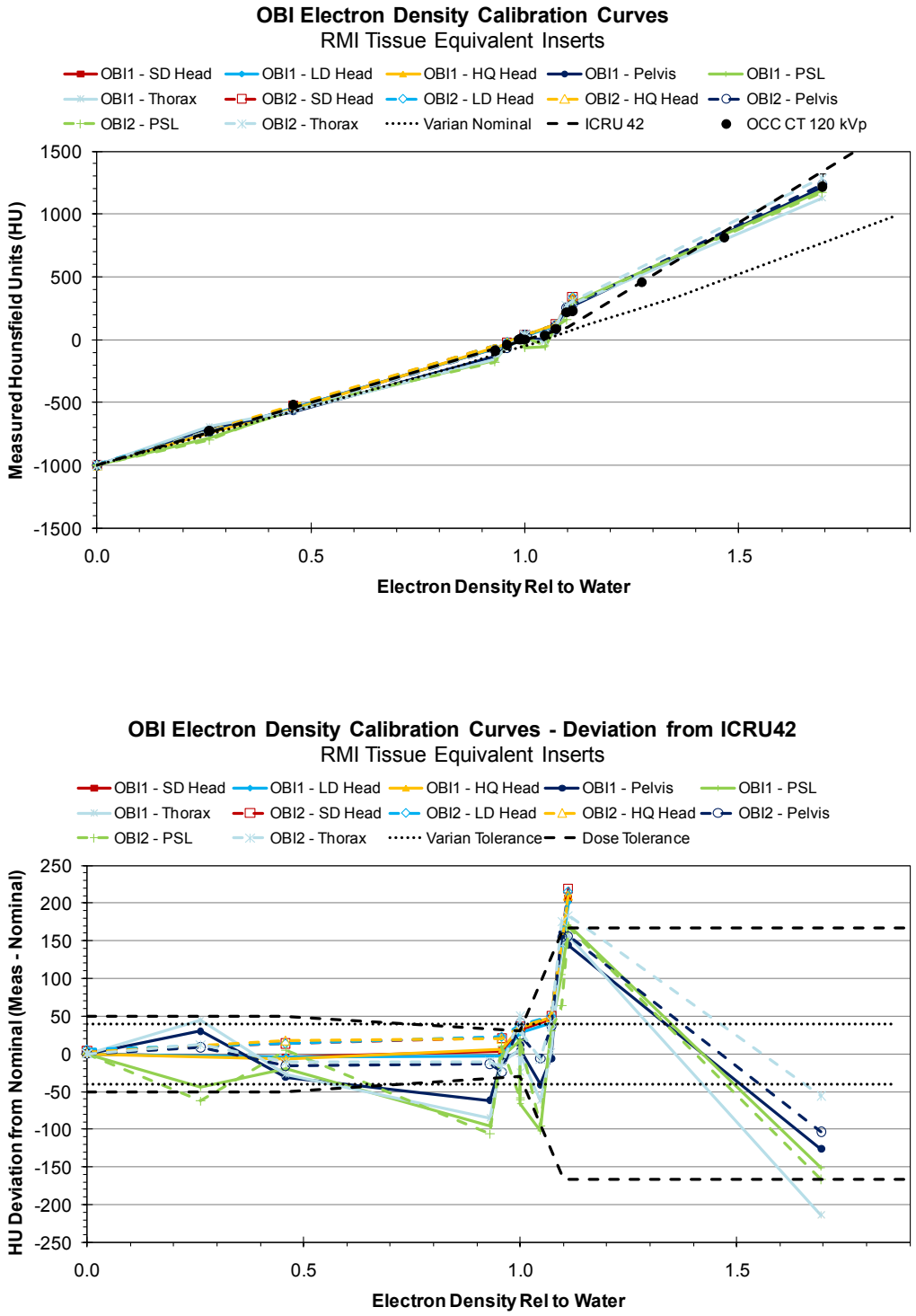


Figure 4.29. – Electron density calibration curves of all clinical imaging modes of the two OBI units in Oxford, measured using ICRU-44 tissue-equivalent materials. Error bars are ±1 SD of the pixel values in each ROI. Top–Measured calibration curves, with the Varian and ICRU-42 nominal curves and the measured 120 kVp curve for the Oxford radiotherapy CT scanner present as references. Bottom–Deviation of measured curves from ICRU-42 nominal. Abbreviations are as in table 4.4.

4.8.4. Investigation of Effect of Equipment Settings on Electron Density Calibration

Longitudinal field of view can be freely modified at scan-time for all OBI CBCT protocols. In an effort to minimise concomitant patient dose it would seem appropriate to make an effort to reduce this as far as possible to just cover the volume of interest. However, the impact of this on the HU sensitometry curve was uncertain.

Scans of the Catphan were acquired using the SD Head mode for longitudinal FOVs of 16 cm, 10 cm and 8 cm, with the phantom aligned as standard with the beam central-axis passing through the centre of the cylinder, and for 10 cm, 8 cm and 4 cm FOVs with the phantom being offset longitudinally by 3 cm to ensure sufficient scattering into the plane of the 'Alignment & Sensitometry' module from material either side. (Because in the standard arrangement the module is not directly in line with the central axis it becomes closer to the periphery of the cone-beam as the longitudinal FOV is reduced.)

Sensitometry curves for the different FOV configurations are presented relative to the Catphan nominal curve in figure 4.30. There is evidence that both the location of the slice with respect to the edge of the cone-beam, and the length of the longitudinal FOV, significantly influence the sensitometry curve, particularly for higher density materials. Further work is required for a full characterisation, but as a result of this initial investigation the current recommendation in Oxford is to ensure that the longitudinal FOV extends at least 4 cm beyond the clinical volume that it is intended to view on the scan.

A limitation of the OBI system is that there is no interlocking mechanism to ensure the correct bow-tie filter is mounted for any particular scanning mode. Indeed, it is also possible to acquire a scan with no bow-tie present. However, it is extremely easy to forget to check or change the filter so an investigation was performed to determine the dependence of the sensitometry curve on the bow-tie filter. A series of scans were acquired, both for full and half-fan modes, where either the wrong or no filter was mounted. It is clear from figure 4.31 that the correct bow-tie filter is crucial for the HU - electron-density conversion curve to be within tolerance. Indeed, the measured curves, and hence both electron-density and relative contrast, are very considerably outside tolerance when the wrong or no filter is used. This is a serious design fault and a procedure has been introduced in Oxford to minimise the risk of this occurring.

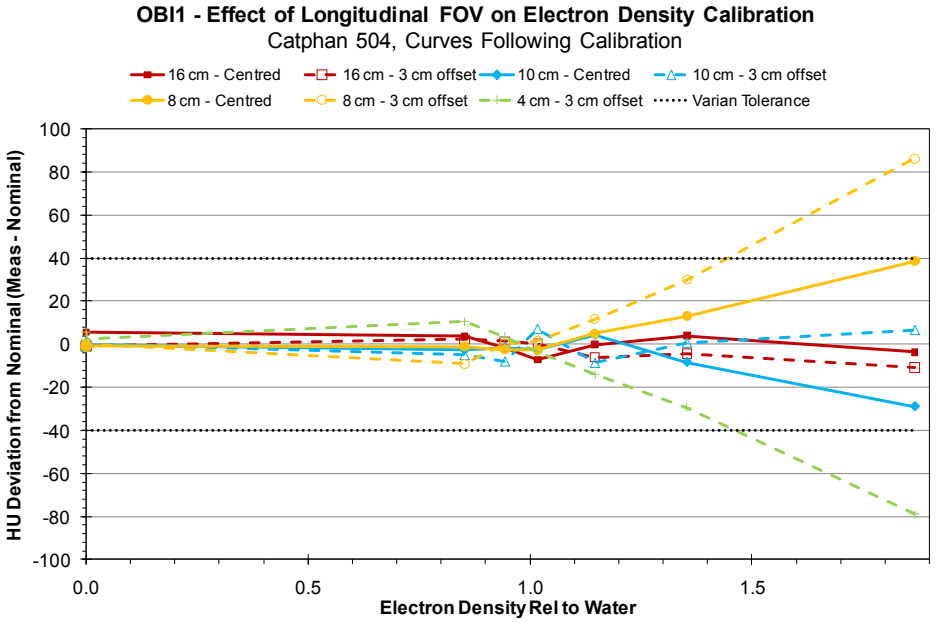


Figure 4.30. – Influence of longitudinal field of view (FOV) on OBI electron density calibration.

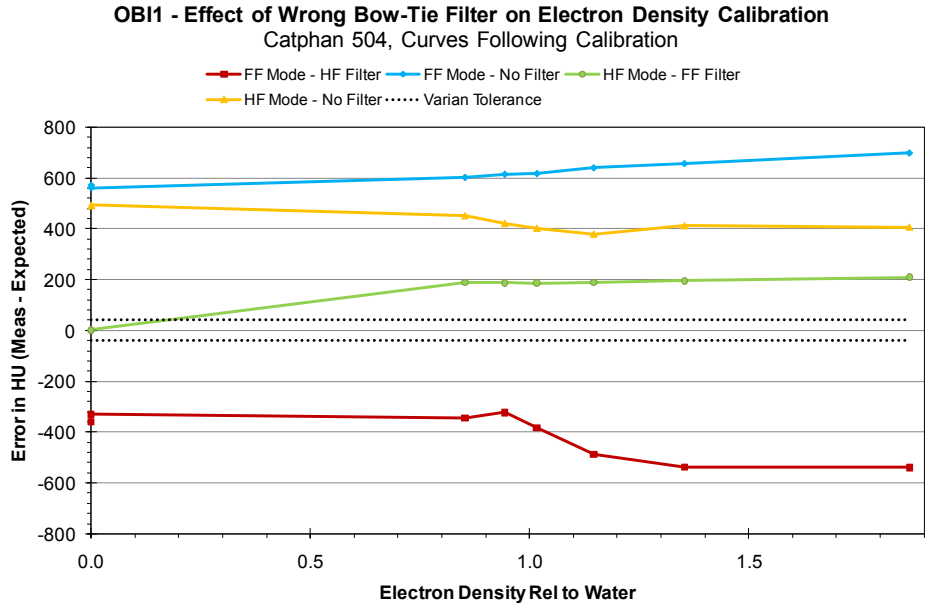


Figure 4.31. – Influence of wrong or missing bow-tie filters on OBI electron density calibration.

4.8.5. Contrast-to-Noise Ratio (CNR)

Over time the radiographers in Oxford have built up considerable experience scanning patients with the conventional CT scanner. It was therefore desirable to draw comparisons between this and the Varian OBI solution. In the following examples detailed measurements were made on one of the new OBI units.

Catphan scans were acquired using each imaging protocol on the OBI, as specified in table 4.4, selecting a slice thickness of 2.5 mm but keeping all other parameters at their default settings (maximum axial and longitudinal FOVs and an image matrix of 384×384). Scans of the phantom were also acquired with the conventional CT scanner using the standard abdomen protocol at FOVs of 25 cm, 50 cm and 65 cm (protocols O5–O7 in table 4.3).

CNR results calculated by IQWorks for acrylic, LDPE and PTFE are shown in figure 4.32 in which it is immediately apparent that each of the conventional CT protocols is superior to any of the OBI ones. Another observation is that for the conventional CT scanner CNR improves with increasing FOV, which is as expected because more photons contributing to the correspondingly larger pixel size will result in less Poisson noise. However, because there is also significant variability in CNR between materials the scaling of the chart makes it difficult to interpret the results further. Therefore, a second chart is included where the results for each material are normalised with respect to that for the 25 cm CT scanner scan.

An interesting result is that across all materials the relative CNR of the various imaging protocols is approximately the same. Also, the CBCT protocols roughly fall into two groups of similar CNR: Low and Standard-Dose Head, and all other protocols. It is also noted that there is a considerable difference between High-Quality Head and Standard-Dose Head (due to dose reduction by a factor of 5, from 21.6 mGy to 4.3 mGy) but almost negligible difference between Standard-Dose Head and Low-Dose Head (where there is dose reduction by a factor of 1.5, from 4.3 mGy to 2.2).

To investigate these results further it is necessary to highlight the dependency on dose. According to Poisson statistics, the CNR should be proportional to the square root of the dose d_p in any pixel. i.e.

$$CNR \propto \sqrt{d_p} \tag{4.9}$$

Now, the dose per pixel is proportional to both the $\mathrm{CTDI}_{\mathrm{vol}}$ for the slice and

the pixel area a^2 where a is the dimension of the side of a square pixel.

$$d_p \propto \mathrm{CTDI}_{\mathrm{vol}} a^2 \tag{4.10}$$

Therefore,

$$CNR \propto \sqrt{\mathrm{CTDI}_{\mathrm{vol}}}\, a \tag{4.11}$$

and hence

$$\frac{CNR}{\sqrt{\mathrm{CTDI}_{\mathrm{vol}}}\, a} = \text{constant} \tag{4.12}$$

Figure 4.33 charts the quantity in equation 4.12 for each material and acquisition mode considered above, with the data points for each material being normalised to the result for the 25 cm FOV CT scan. In the discussions of CNR results earlier in this chapter it was noted a number of times that there may be an intrinsic uncertainty in CNR measurements due to a potentially large contrast value being divided by a numerically small noise value, such that even small fluctuations in the noise value were manifested in artificially large changes in CNR. When discussing the time-trend data for the Catphan CNR measurements in figure 4.21 it was noted that the measured values fluctuated within an envelope of ±20% of the baseline. This was taken as a 'worst case' estimate of the uncertainty in the CNR calculation in the current experiment and used for the error bars on figure 4.33.

A number of interesting results are revealed by this figure. Firstly, the results are similar across all materials, which is to be expected for materials with smoothly varying linear attenuation coefficients over the 10–100 keV range, and against which the individual scan modes have been calibrated. i.e. protocol kVp should have little impact on contrast, both because it has been calibrated out and because these are 'well-behaved' engineering materials. Also, the results for the 25 and 50 cm CT scanner protocols are almost identical, which is exactly in line with expectation. However, the results for the 65 cm 'wide FOV' CT scanner mode are considerably larger than the other two. Although this could be just on the limits of the estimated experimental uncertainty it is unlikely because the value is consistently a factor of ~1.6 times that of the other modes, with very little fluctuation between materials. This result implies that the 'wide FOV' mode achieves improved CNR at any given dose point, even accounting for pixel size, which is somewhat counter-intuitive given that clinical users report extrapolation artefacts in patient images acquired under this mode. However, a limitation of this experiment is that the Catphan is of a relatively small diameter (20 cm) and is positioned at the centre of the

FOV, so that the reconstruction of its image is from a fully sampled set of projections. Given that the 'wide FOV' reconstruction algorithm attempts to minimise an expectation function when estimating the 'missing' data required to fill the extended FOV it is likely there are noise reduction steps in place, possibly to reduce the artefacts introduced at the periphery of the FOV. It is therefore thought that the perceived improvement in relative CNR may not be representative and that a more detailed investigation of the effect on other metrics is required.

Another interesting result is that, within the uncertainty limits, all CBCT modes except Pelvis perform the same. This is reassuring because it indicates the system is limited by Poisson statistics rather than data processing steps. However, it is unclear why the performance of the Pelvis mode is significantly poorer than the rest. Indeed, with reference to table 4.4, at a $CTDI_{vol}$ of 19.6 mGy the Pelvis mode is one of the highest dose modes, so the fact that it appears to be underperforming merits further investigation. Although the metric calculated for the LDPE test object on the Low Dose Head mode appears to be superior to those of other modes it is thought this is still within the experimental uncertainty because, in contrast to the 'wide FOV' case discussed above, this value varies within its error bar between the acrylic and PTFE materials.

In a further experiment the phantom was scanned again using different matrix sizes (384×384 and 512×512) and fields of view (25 cm – which included 25.5 and 25.6 cm, and 45 cm). The results for the acrylic insert are presented in figure 4.34. As expected, there is always a boost in CNR when reconstructing at the smaller matrix size, and this is of the order of 30%. Furthermore, it is interesting that the Pelvis protocol, when reconstructing at its smallest 25.6 cm FOV, has comparable CNR to the Pelvis Spot-Light and High-Quality Head protocols. This suggests there is no CNR penalty from using a half-fan mode for relatively small FOV scanning. Given that FOVs smaller than 25 cm are unlikely to be required in the majority of clinical scenarios, and putting concomitant dose considerations aside, the Pelvis protocol may be appropriate as a general purpose option which can be utilised in most situations. This would have the advantage of avoiding the necessity to regularly change the bow-tie filter, but the more time-consuming gantry rotation over 360° instead of 200° may have clinical implications. There may also be patient dose implications due to the X-ray tube irradiating from all angles, but a detailed consideration would involve modelling doses to organs within the irradiated volume and is outside the scope of this study[71].

To verify whether these results were also in line with expectation the metric in equation 4.12 was calculated and is plotted in figure 4.35, this time normalised to the SD Head result for the 25 cm, 384 $\times$ 384 case. For the same reasons as discussed above, the error bars are again set at ±20%. This time all modes perform similarly, within the limits of experimental uncertainty, except for the large FOV Pelvis scans, which again underperform. Again, further investigation of this is required.

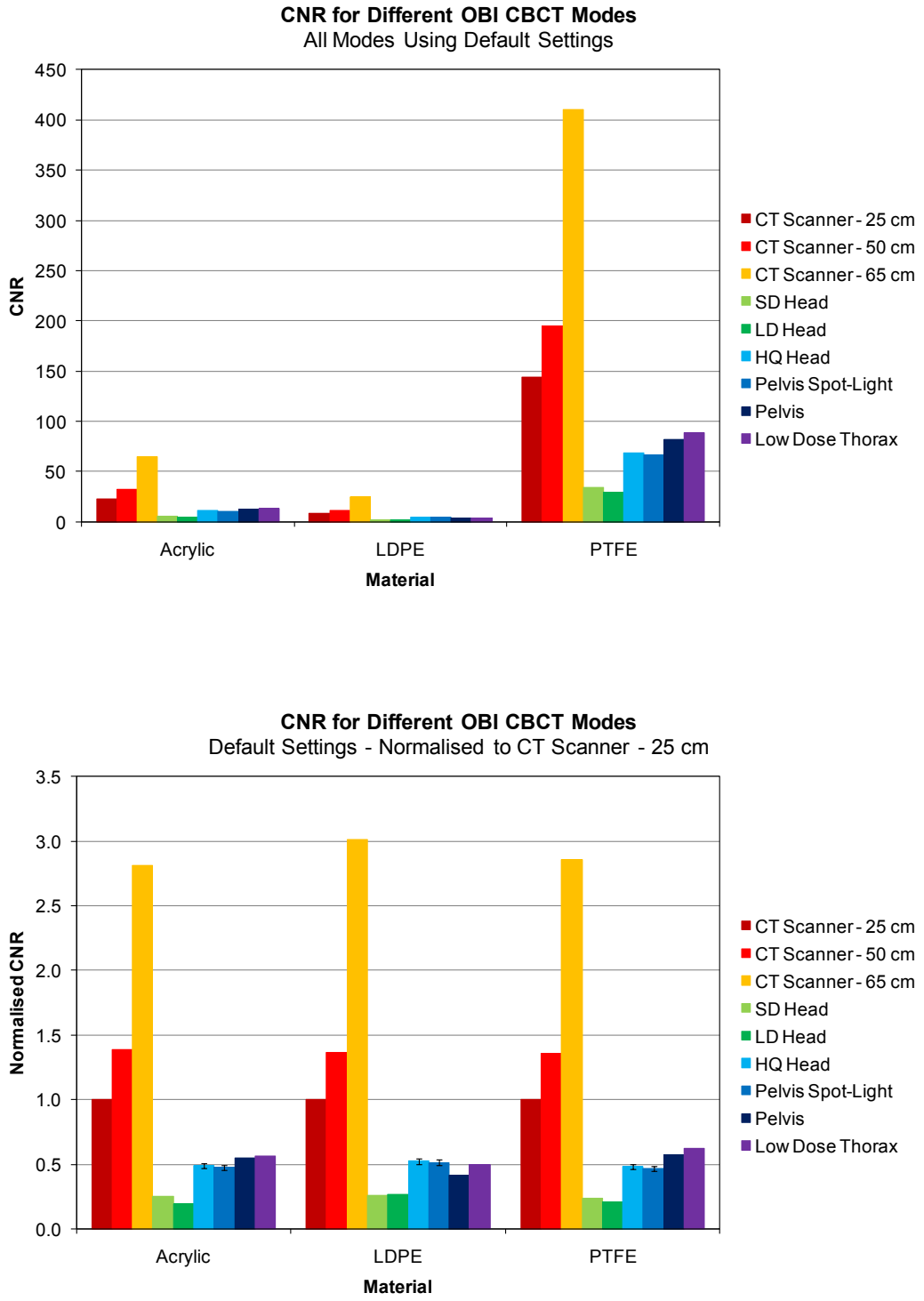


Figure 4.32. – Contrast-to-noise ratio (CNR) of default OBI acquisition modes relative to standard modes of the radiotherapy CT scanner in Oxford Cancer Centre (OCC). Error bars are indicative of ±1 SD in the pixel values used in the CNR calculation and are included for illustrative purposes only because there is significant uncertainty in the calculation of these. Top–Absolute CNR for three reference materials. Bottom–Curves normalised to the CNR of the 25 cm CT scanner result. Abbreviations are as in table 4.4.

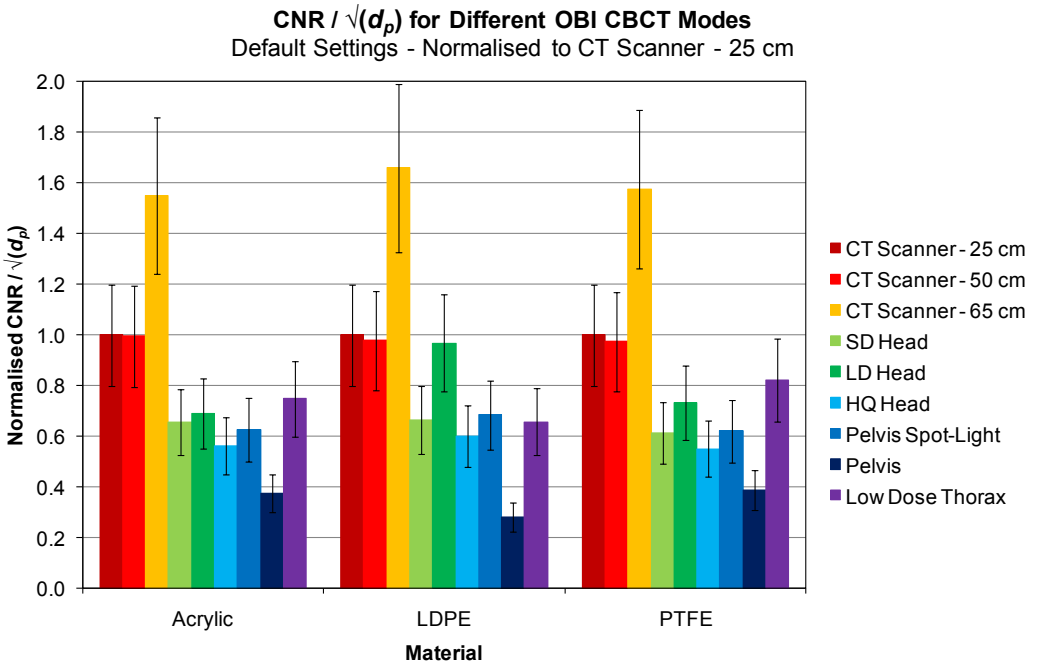


Figure 4.33. – Ratio of contrast-to-noise ratio (CNR) and square root of dose per pixel of default OBI acquisition modes, normalised to the performance of the Oxford CT scanner under protocol O5 in table 4.3. CT protocols O6 and O7 have also been included for comparative purposes. Error bars represent ±20% and are explained in the main text. Abbreviations are as in table 4.4.

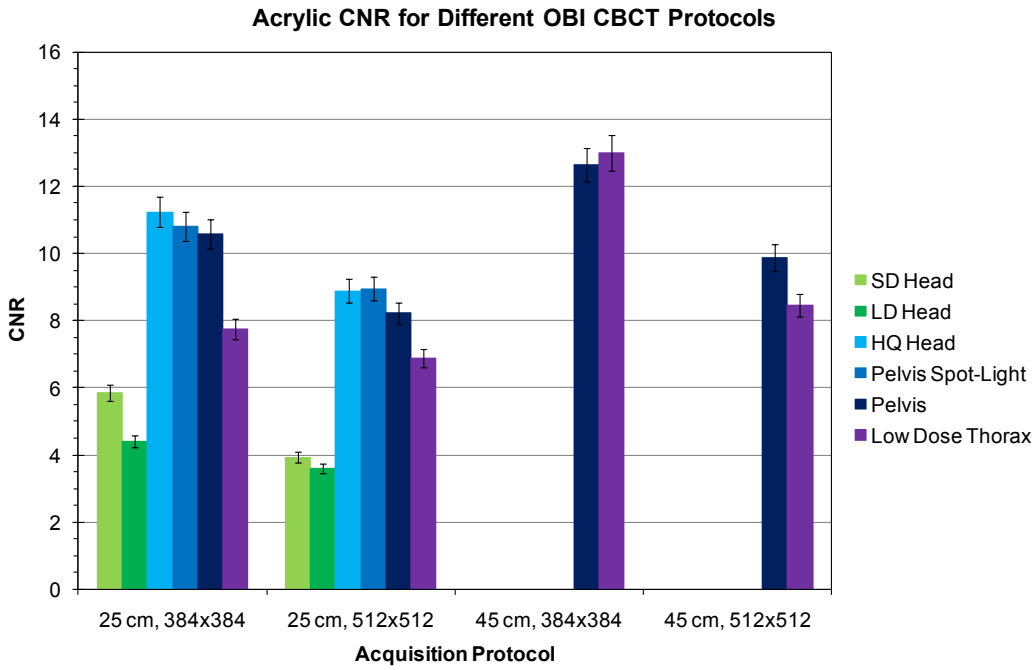


Figure 4.34. – Absolute contrast-to-noise ratio (CNR) of acrylic against polystyrene for different OBI acquisition settings and protocols. Error bars are indicative of ±1 SD in the pixel values used in the CNR calculation. Abbreviations are as in table 4.4.

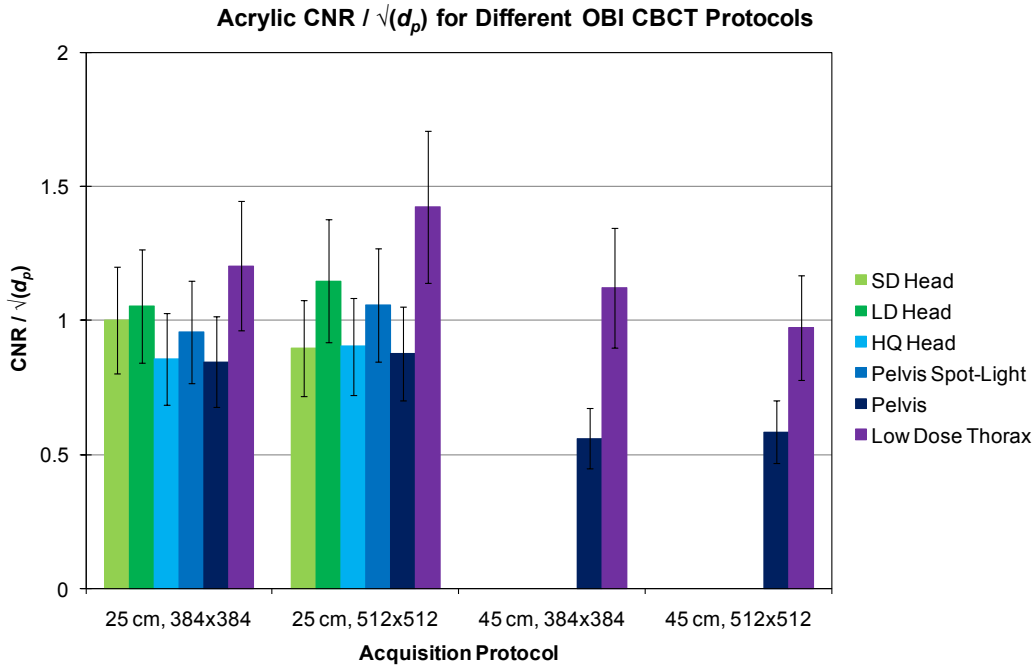


Figure 4.35. – Ratio of contrast-to-noise ratio (CNR) to square root of dose per pixel for different OBI acquisition settings and protocols, normalised to the result for the SD Head protocol at 25 cm, 384 × 384. Error bars indicate ±20% and are explained in the main text.. Abbreviations are as in table 4.4.

	f_{50} (cycles/mm)		
Mode	512×512	384×384	$\frac{512\times512}{384\times384}$
SD Head - 25 cm FOV	0.54	0.42	1.3
Pelvis - 26 cm FOV	0.56	0.51	1.1
Pelvis - 45 cm FOV	0.41	0.35	1.2

Table 4.6. – MTF f_{50} results for different acquisition modes and reconstruction settings

4.8.6. Modulation Transfer Function (MTF)

IQWorks was used to analyse one of the MTF beads in the same series of scans as acquired for the CNR experiments discussed in the previous section. Results for the Standard-Dose Head and Pelvis protocols are plotted in figure 4.36, with comparative results also included for standard, clinically used head and pelvis protocols of the conventional CT scanner.

Interestingly, the MTF curves for the OBI system are generally better than those of the conventional CT scanner. Although this is somewhat surprising it is noted that the CT scanner protocols utilised do not employ any form of edge enhancement, so this is not a true indication of relative performance under fully optimised conditions[45, 175]. Nevertheless, these results are representative of the performance of the two systems for standard modes in current clinical use. Another important result is that there is a considerable boost in MTF from reconstructing at a matrix of 512 × 512 rather then 384 × 384. This is presented numerically in table 4.6, where it is evident that the improvement in f_{50} when moving to the larger matrix size is 1.2±0.1 across all modes. This is consistent with the factor of $512/384 = 1.3$ one would expect from the smaller pixel size.

Serious consideration must therefore be given to the balance between CNR and MTF when selecting a matrix size appropriate for a particular clinical application.

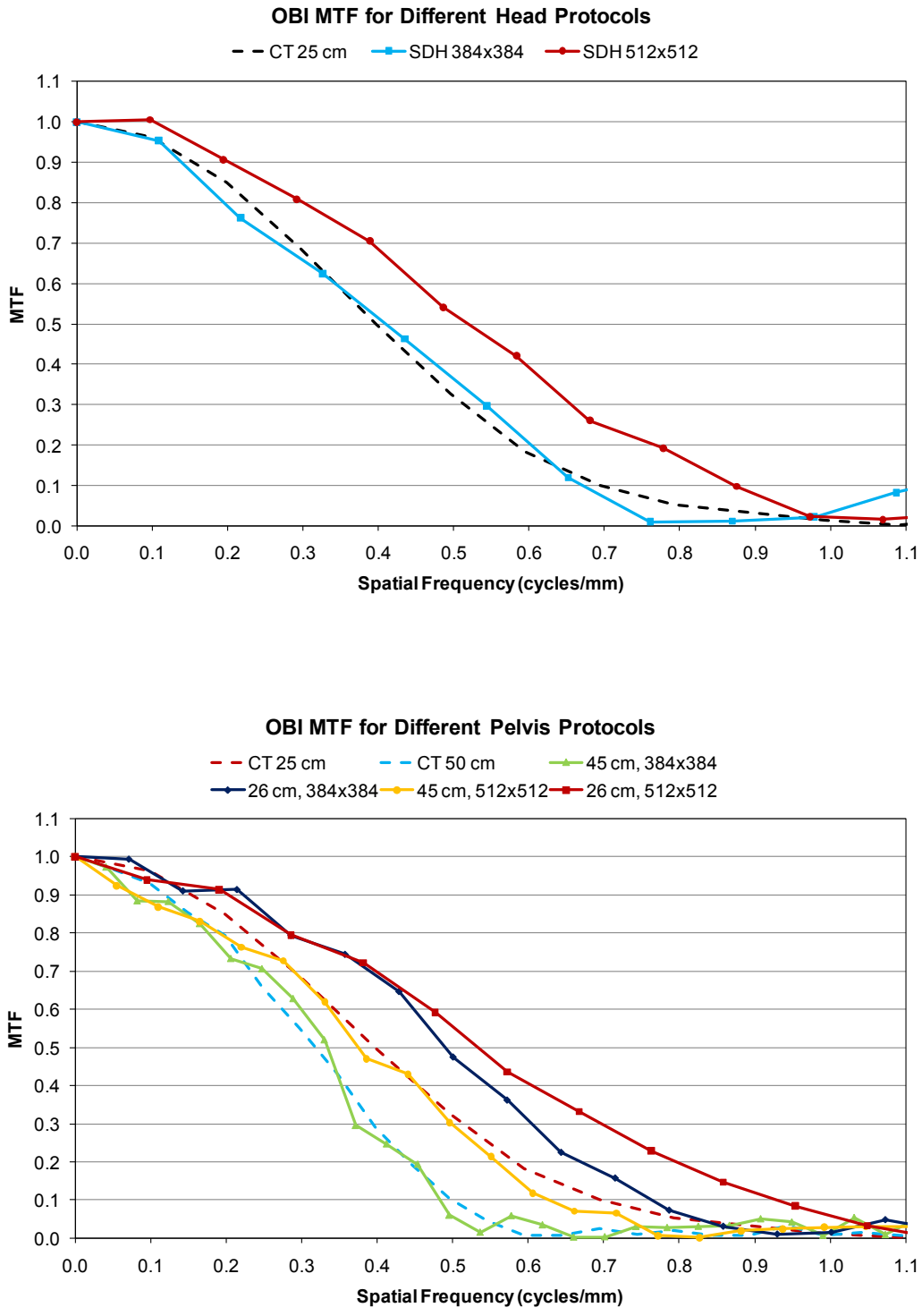


Figure 4.36. – Modulation transfer function (MTF) curves of OBI standard dose head (SDH—top) and pelvis (bottom) protocols, in comparison with those of the radiotherapy CT scanner in Oxford Cancer Centre (OCC)

4.8.7. Uniformity

By visual inspection, CBCT images generally appear noisier and less uniform than those of the conventional CT scanner (compare, for example, the images in 4.37). The subjective perception of higher noise is consistent with the objective CNR measurements reported above. However, a detailed numerical analysis is required to confirm the visual observation for uniformity.

A known issue with the OBI system is that full-fan modes suffer from a 'crescent' artefact caused by the bow-tie filter slipping relative to the focal spot of the X-ray tube as the gantry rotates. This is clearly visible in the image in figure 4.37. However, it was discovered that the U1 and 'maximum difference' recommended metrics[156, 345] are rather insensitive to this, essentially because the ROIs inspected in the analysis lie either side – but not on – the artefact.

Instead, the artefact is clearly revealed by a diagonal profile, as illustrated in figure 4.38. Another interesting result from this figure is that, whereas the response of the Edinburgh and Oxford CT scanners is flat across the entire FOV, the CBCT modes tend to fall-off towards the periphery.

The reduced uniformity metrics defined in section 4.7.3 are compared across all OBI imaging modes for both inner and outer FOVs in figure 4.39. These are described in table 4.4. Results from the Edinburgh and Oxford CT scanners (protocols E13 and O5 in tables 4.2 and 4.3) and the Acuity (table 4.5) are included for comparison. In this figure, the larger the value of the uniformity metric the poorer the uniformity. As expected, all uniformity metrics for both CT scanners are better than those of any CBCT imaging protocol. Interestingly, there is significant variability in uniformity between the various CBCT modes and it is unclear why there is a considerable difference between that of the High-Quality and other head modes, when these modes should be identical except for the detector dose per projection. Both integral uniformity and coefficient of variation were sensitive to the CBCT crescent artefact, but the differential uniformity metric was relatively unaffected by it.

These results reinforce the assertion that more than one metric should always be utilised whenever characterising uniformity.

Finally, the influence of FOV and matrix size was investigated for the pelvis scanning protocol, with the results presented in figure 4.40. All metrics are normalised with respect to their values at a 26 cm FOV with a 512×512 image matrix. Although uniformity appears to improve with decreasing matrix size and increasing FOV, this may be due simply to a larger pixel size smoothing out any non-uniformities and further investigation is necessary.

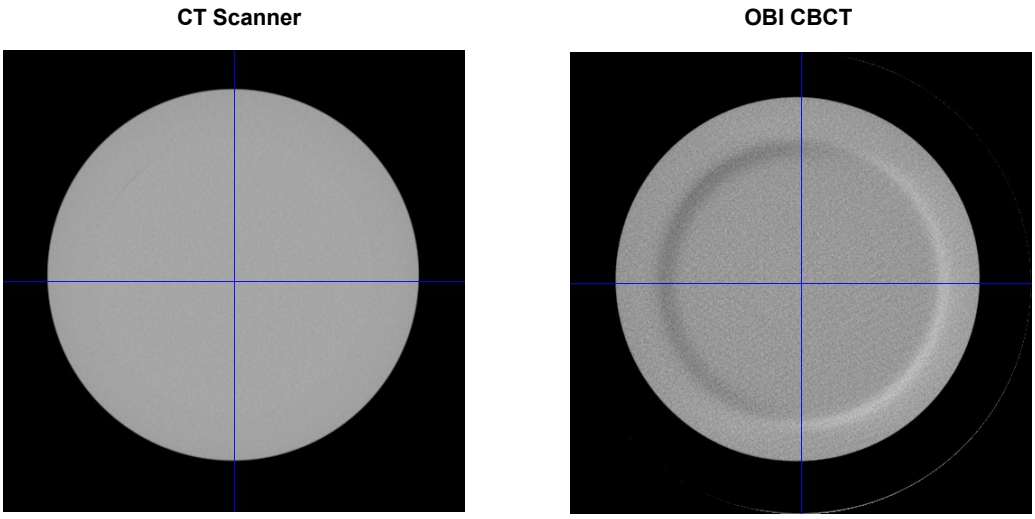


Figure 4.37. – Catphan 504 Uniformity module images acquired using a conventional CT scanner and OBI CBCT.

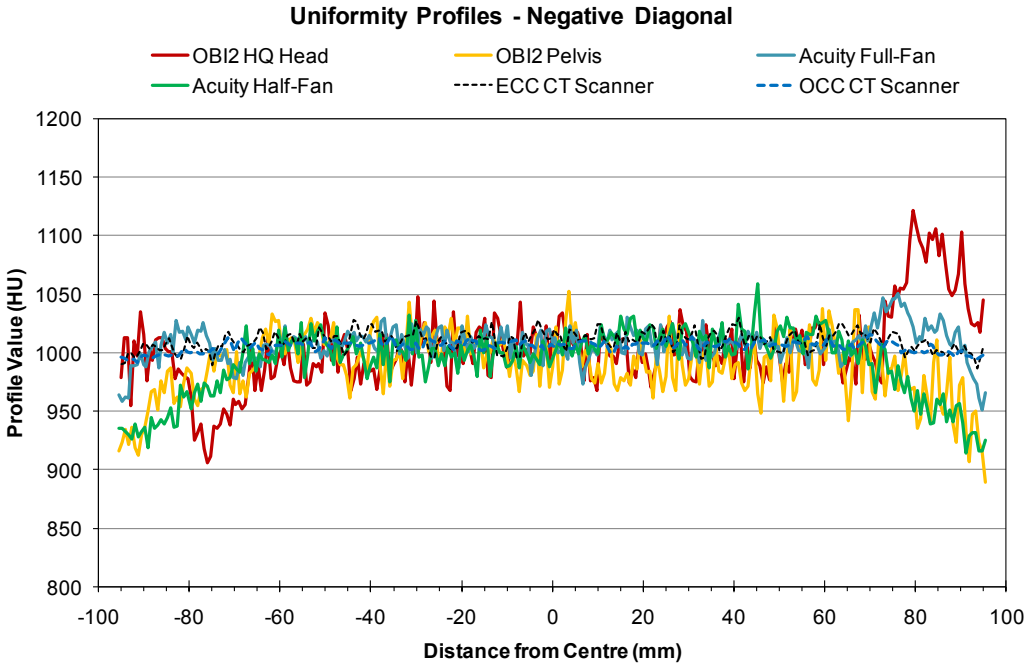
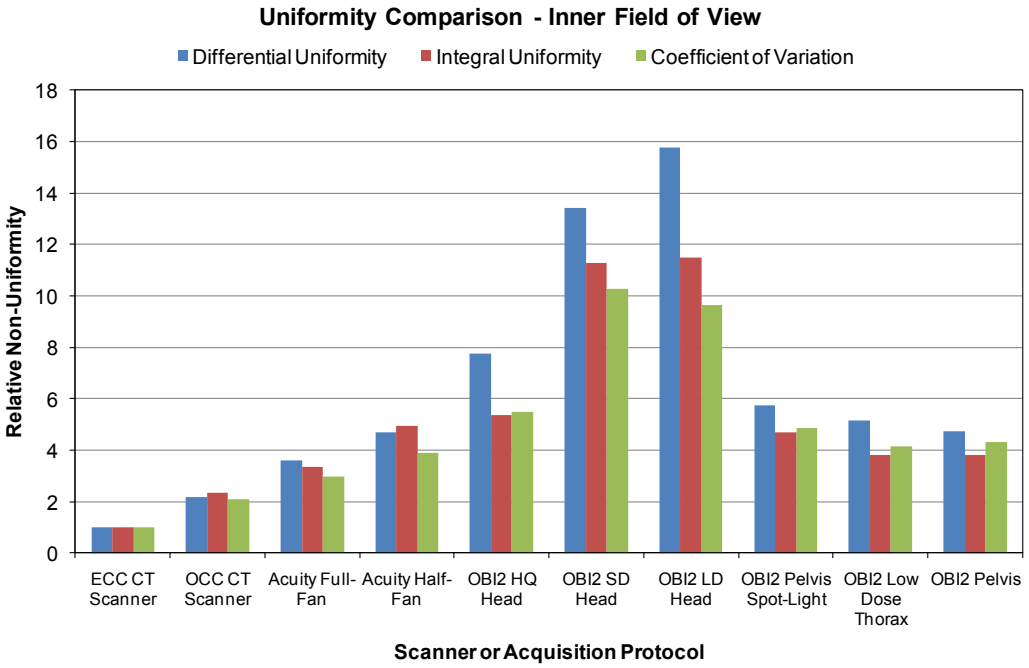


Figure 4.38. – Uniformity profiles of Oxford OBI and Acuity units over the outer field of view of the Catphan 504 uniformity module for full and half-fan acquisition modes. Results from the Edinburgh Cancer Centre (ECC) and Oxford Cancer Centre (OCC) CT scanners are included for the purposes of comparison.



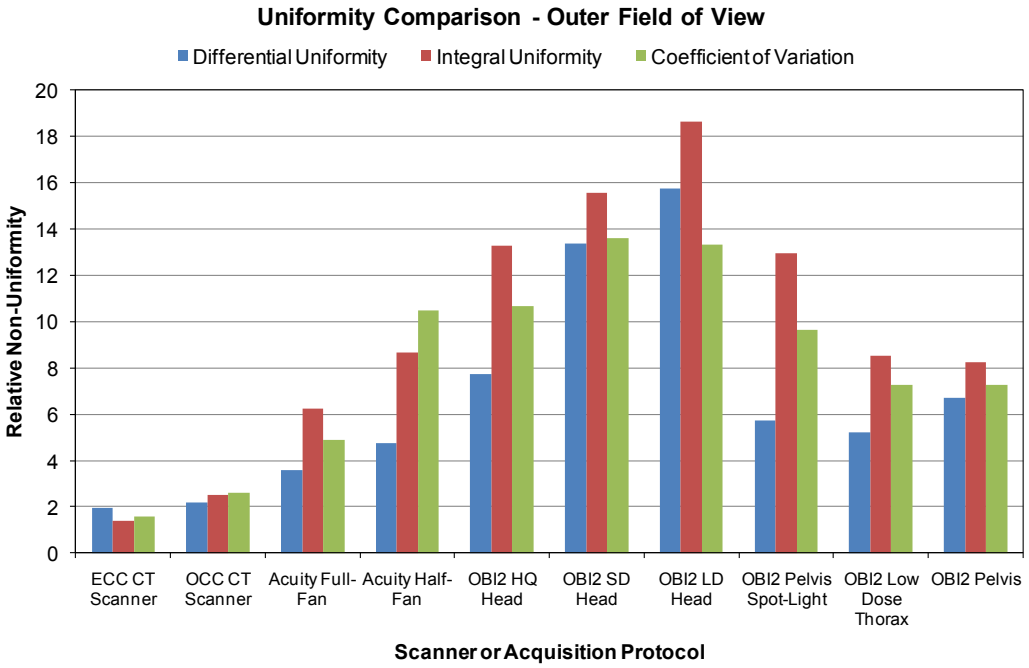


Figure 4.39. – Comparison of worst-case uniformity metrics across all default imaging modes of the OBI and Acuity units in Oxford. Top–Inner field of view (±10 cm). Bottom–Outer field of view (±19 cm). Results for each metric are normalised to those for the inner field of view of the Edinburgh Cancer Centre CT scanner. OBI mode abbreviations are as in table 4.4.

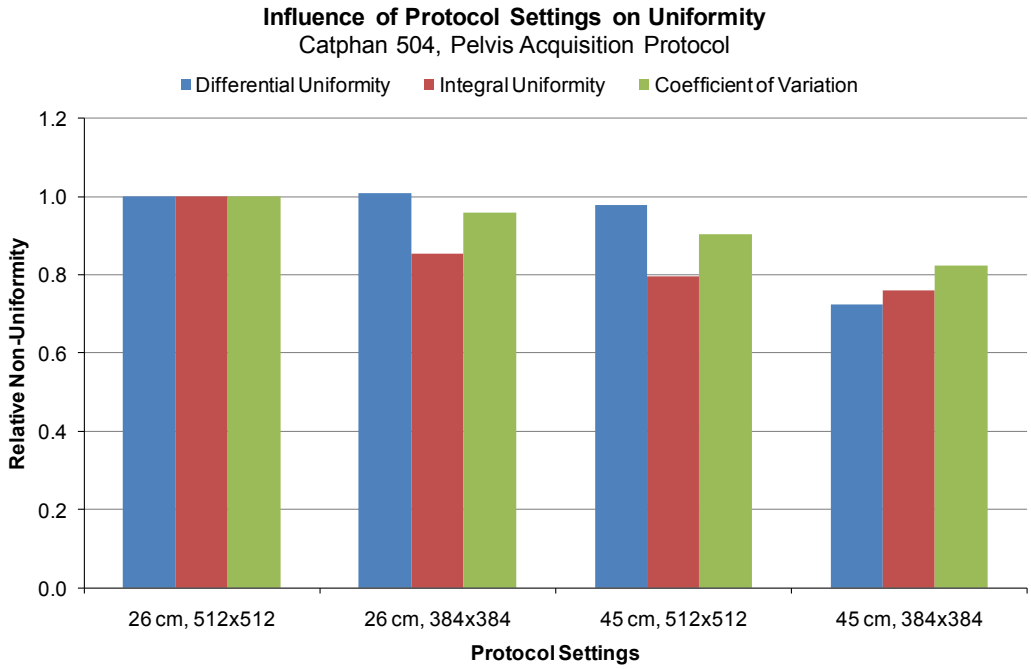


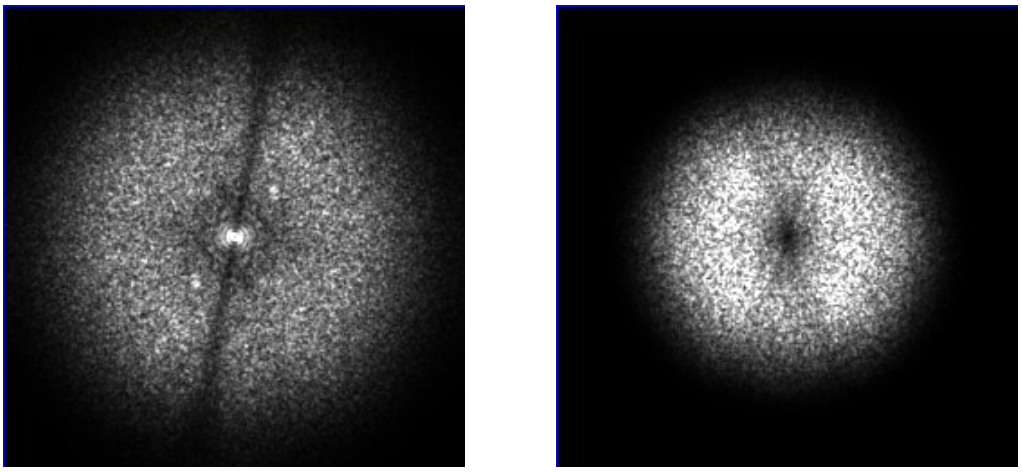
Figure 4.40. – Influence of different OBI pelvis acquisition protocol settings on outer field-of-view uniformity metrics.

4.8.8. Normalised Noise Power Spectrum (NNPS)

Noise power spectral analysis was performed on a High-Quality Head scan of the Catphan uniformity module. A single 256 × 256 ROI was placed at the centre of FOV and 17 adjacent slices were considered to calculate the NNPS. The results for stochastic noise are compared against those for the clinically used head protocol () of the conventional CT scanner in figure 4.41.

From the figure it is apparent that the NNPS of the CBCT scan is approximately two orders of magnitude greater than that of the conventional CT scan, which is consistent with the poor CNR results discussed above. Furthermore, the 2D visualisation of NNPS reveals off-axis structure in the CBCT spectrum which is not present in conventional CT. There also appears to be a rotation of the spectrum so that its axis of symmetry is not the Y axis. This may be due to the CBCT scan being acquired over a partial arc whereas the conventional CT scan employs a full 360° rotation.

Clearly, a great deal more investigation is required before any conclusions can be drawn but this example demonstrates the potential of the IQWorks evaluation framework for performing detailed experiments to determine the root of noise-related issues. Nevertheless, it is reassuring that the results are consistent with those of other researchers[30].



OBI CBCT
High Quality Head Protocol

OCC CT Scanner
120 kVp Head Protocol

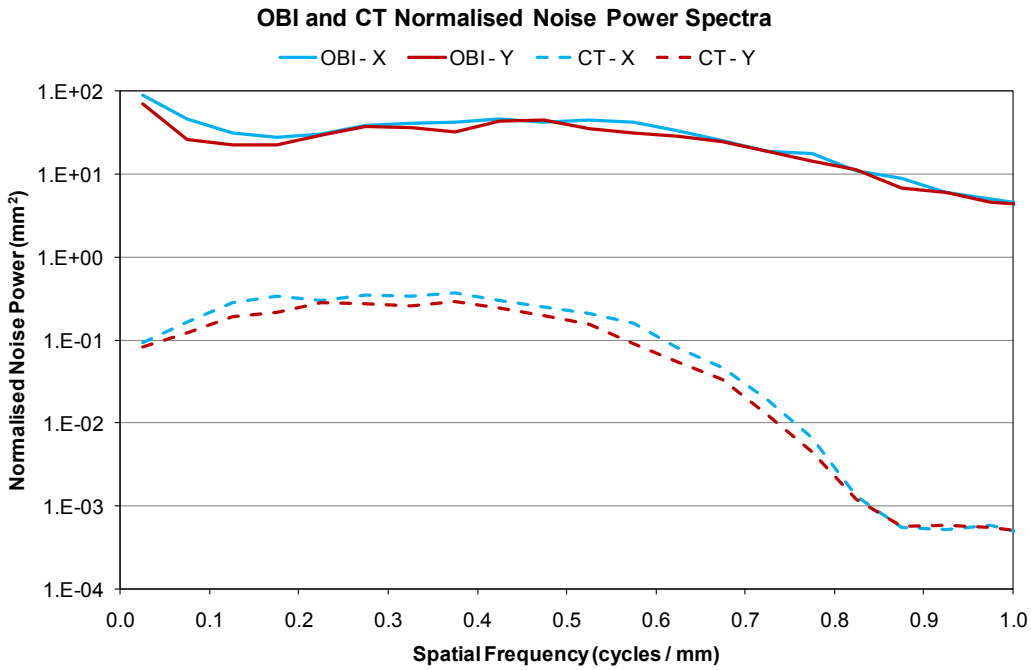


Figure 4.41. – Comparison of OBI and CT scanner normalised noise power spectra (NNPS). Top–2D NNPS for equivalent head protocols of each system. Bottom–X and Y NNPS curves.

4.8.9. OBI Time-Trends

During the fortnight over which the first OBI unit was commissioned a daily CT scan was acquired of the Catphan 504 phantom using the Standard Dose Head protocol, described in table 4.4, with a slice thickness of 2.5 mm, axial FOV 25.5 cm, longitudinal FOV 17 cm and image matrix of 512×512. Follow up scans were then performed during regular QA sessions.

Sensitometry trends for the LDPE, acrylic and PTFE inserts are plotted in figure 4.42, in which the day to day fluctuations for all materials are within an envelope of ±30 HU about the baseline. Although these results exhibit significantly greater variability those reported above for the Edinburgh CT scanner (±7 HU) they are still well within the ±40 HU Varian tolerance and there is no evidence of systematic drift. Of all materials, the highest density PTFE appears to vary about the mean the most.

CNR time-trend results are presented in figure 4.43. Over the first four months these are of the same order of magnitude as the Edinburgh CT scanner, being within ±20% of the baseline. However, over the last two months there appears to be a systematic drift towards higher CNR, with it currently lying just below 20% above baseline for all materials. Given that figure 4.42 demonstrates that the HU calibration curve has not changed over this period it appears that the overall noise level in the system is decreasing. This could be indicative of a change in X-ray tube output, so warrants further investigation.

In addition, geometric linearity is also very stable, as illustrated in figure 4.44, with the results agreeing well with those observed for the Edinburgh CT scanner: within ±1.5% of baseline for all measurements. For inter-hole distances this corresponds to an absolute uncertainty of ±0.75 mm, which is well within the 2.0 mm tolerance for OBI measurement accuracy[185].

Employing the IQWorks evaluation framework to underpin the OBI QA programme has clearly demonstrated benefits. Regular storage of detailed, quantitative results to database both facilitates the identification of time-trends and enables deeper investigations to be performed if required.

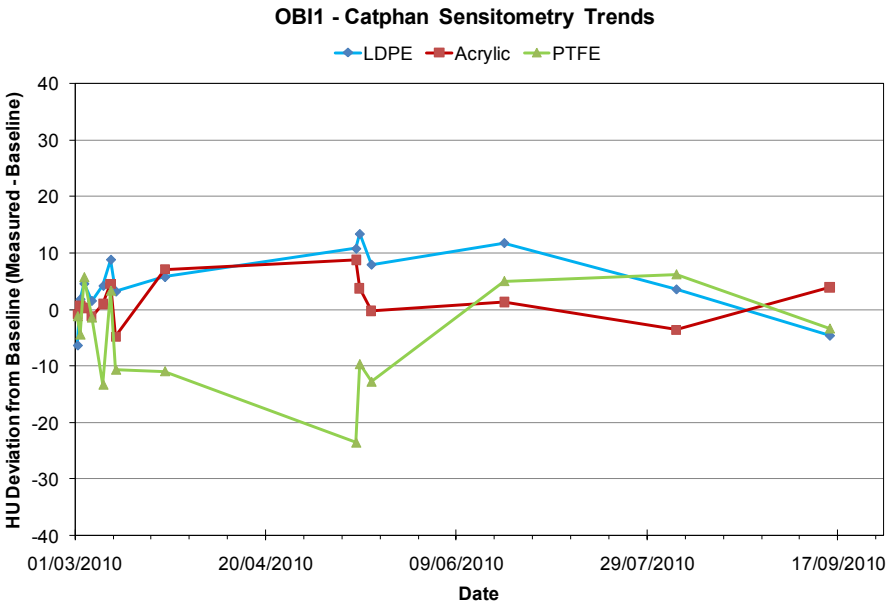


Figure 4.42. – Sensitometry time-trends for OBI unit V1 in Oxford. Data are presented for the sensitometry materials in the Catphan 504 phantom scanned using the Standard Dose Head protocol.

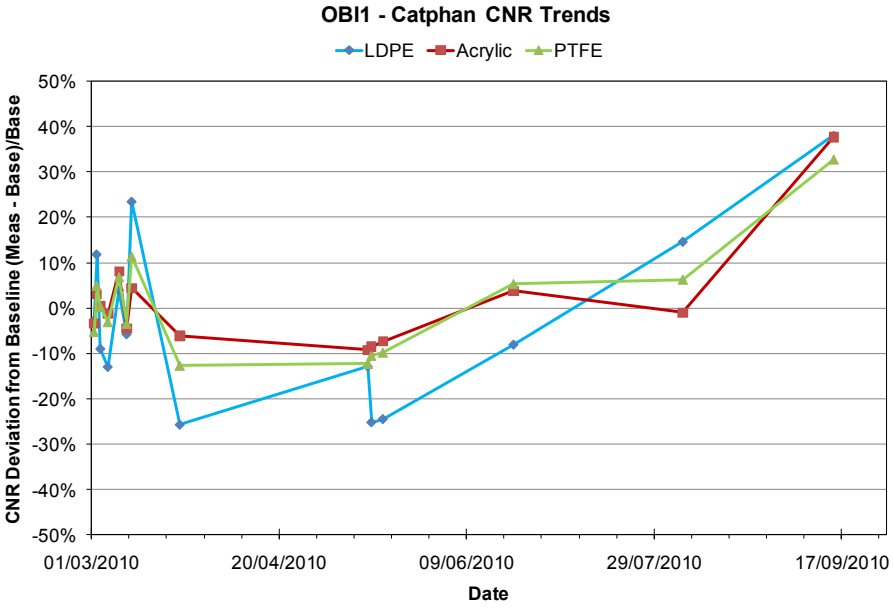


Figure 4.43. – Contrast-to-noise ratio (CNR) time-trends for one OBI unit in Oxford. Data are presented for the sensitometry materials in the Catphan 504 phantom scanned using the Standard Dose Head protocol.

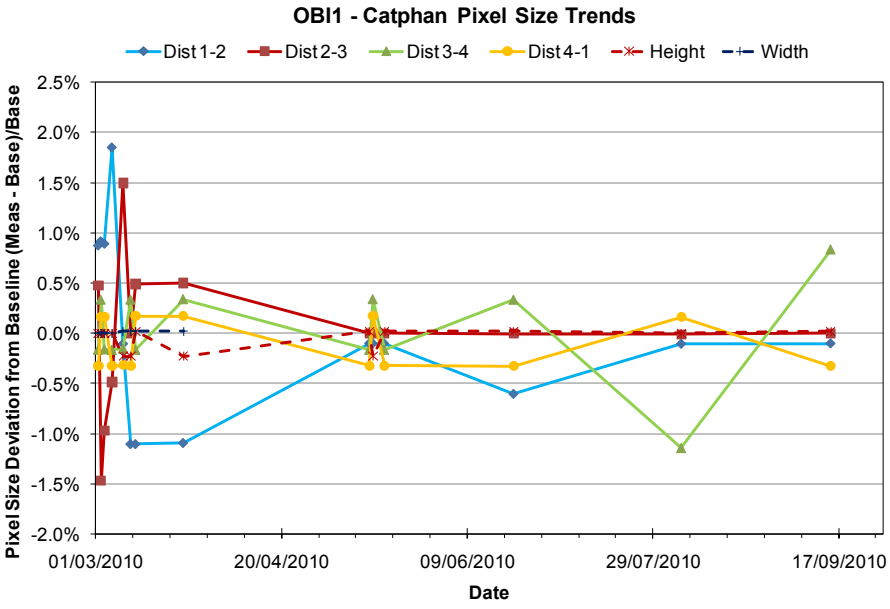


Figure 4.44. – Pixel size time-trends for one OBI unit in Oxford. Data are presented for the four distance measurements and the width and height of the Catphan 504 phantom scanned using the Standard Dose Head protocol.

4.9. Conclusion

Numerous examples have been presented of the IQWorks evaluation framework being successfully utilised to analyse a range of established phantoms for the performance evaluation of CT and CBCT. Applying a consistent approach across all phantoms and imaging units enables intercomparisons of both equipment and techniques. In addition, the flexibility of IQWorks analysis trees permits existing analysis schemes to be modified *ad hoc* to accommodate specific experimental requirements or improvised new phantoms. This can be achieved whilst still maintaining compatibility with other phantoms and schemes at the algorithmic level.

Two phantoms containing materials with attenuation properties similar to that of real tissue were utilised to examine the electron density calibration curves of two GE CT scanners. It was confirmed that the 120 kVp and 140 kVp beams of these units yield curves sufficiently close to the ICRU 42 standard curve that dose calculations would be accurate to within 2% if the ICRU curve was used in the treatment planning system, rather than curves tailored to the individual scanners. This result was consistent with Battista and Bronskill's formalism, with the effective energies $\tilde{E}$ of the 120 and 140 kVp beams falling within the specified 60 kVp $< \tilde{E} <$80 kVp range of applicability. Furthermore, in line with expectation it was confirmed that the electron density calibration curve is sensitive only to kVp and not other scanner acquisition settings.

Regular quality assurance measurements on both CT scanners demonstrated stable performance over an extended time frame, with the mean pixel values of reference materials within ±5 HU of baseline for soft-tissue and within ±12 HU for bone at all times for a phantom positioned on the flat couch-top, but within the tighter ranges of ±1.5 HU for soft-tissue and ±5 HU for bone when a phantom was mounted on a jig hanging from the couch. It was concluded that phantom placement influences the reproducibility of sensitometry measurements which is important to be aware of because this is representative of real patient scanning scenarios.

Modulation transfer function (MTF) measurements were consistent with those of other researchers and MTF f_{50} was found to lie consistently within ±0.06 cycles / mm of baseline. It was noted that care should be taken when undertaking MTF measurements to model the scenario being investigated: an experiment to characterise intrinsic detector performance may not yield results representative of the resolving power of a real clinical protocol. In addition, in all experiments it was found that MTF improves as the field of view is reduced

and it is suggested that best practice is to make use of the smallest FOV which will comfortably accommodate a patient, rather than rely on default protocol values.

Contrast to noise ratio (CNR) results were found to always fall within ±10% of baseline over an extended time period. In addition, an interesting result was that for higher dose protocols, where the image noise is lower, the variability in CNR over time appears to be greater. It is thought this is due to the CNR calculation being extremely sensitive to small changes in noise when the numerical value of the noise quantity on the denominator is itself small. Although this means CNR calculations will be sensitive to small changes in system performance – as was demonstrated in the Edinburgh QA programme when the X-ray tube was "spitting" due to an imperfect vacuum – it is essential a baseline level of expected fluctuation is established so that false positives can be discounted.

An important conclusion arrived at a number of times in the discussion above is that dedicated phantoms containing biologically representative tissue-equivalent materials are necessary for the characterisation and verification of the HU - electron-density conversion curve. Unfortunately, despite its comprehensive features for general purpose imaging evaluations the Catphan does not appear to be suitable for this application. Furthermore, the limited field of view of the Catphan inhibits its use in the optimisation of many radiotherapy imaging scenarios, especially the most challenging cases where a patient may be positioned towards the edge of a large field of view. There is a tangible need for a new comprehensive image quality phantom, similar in principle to the Catphan, but covering a larger field of view and containing tissue-equivalent sensitometry inserts. This is a current area of research by the author.

Another important finding was that it is important to routinely calculate at least two uniformity metrics, with the standard metric for CT uniformity as recommended in IPEM Report 32[156] itself being insensitive to significant changes detected by other metrics. The problem is that some whilst some metrics are localised, others assess global performance. When considering the overall performance of a system it is therefore essential to choose appropriate metrics which cover likely eventualities.

During the commissioning of two OBI units in Oxford the IQWorks framework clearly demonstrated its worth. Being able to quickly and automatically undertake repetitive, time-consuming analyses enabled detailed characterisations of clinical imaging modes which would not otherwise have been possible. Furthermore, key performance factors were identified which must be con-

sidered when optimising scan parameters for individual patient scenarios. A significant boost in MTF f_{50} of 20-30% is achievable by moving from a 384×384 to 512×512 pixel matrix, although there is a corresponding reduction in CNR of the order of 30% due to the smaller pixel size. Therefore, the balance between MTF and CNR when choosing the image reconstruction matrix requires further consideration and consultation with clinical users. It is expected these will have a real clinical impact.

Measurements suggested that the CT systems considered in this study were all limited by Poisson noise, with noise levels changing as expected with changes in CTDI and pixel matrix. Interestingly, this analysis has revealed what appears to be an underperforming OBI imaging mode, and this is currently the subject of investigation.

As a platform from which to manage a QA programme based on objective image quality metrics, the value of IQWorks has also been clearly established.

Chapter 5.

Application to Other Radiotherapy Imaging Modalities

5.1. Overview

In the previous two chapters detailed examples were presented of the IQWorks image analysis framework being applied to the performance evaluation of EPIDs and X-ray CT. Although these are the two imaging modalities most commonly utilised as part of radiotherapy treatment planning and delivery IQWorks can also be used to assess images from almost any other modality, as long as digital image data can be obtained for input into the software. Furthermore, whilst the emphasis in this work has been on the implementation of objective image quality methodologies into radiotherapy practice the framework is equally applicable to assessment and optimisation exercises in the sphere of diagnostic imaging.

This chapter outlines some examples of using IQWorks to analyse images from a range of other modalities commonly encountered in radiotherapy.

One of the radiographic modes of the OBI PaxScan 4030CB detector is characterised in terms of its modulation transfer function (MTF) and normalised noise power spectrum (NNPS). In addition, the potential application of the new QEPI1 portal imaging phantom, described in chapter 3, to the performance evaluation of kilovoltage energy imaging systems is investigated.

A new phantom is introduced for the geometrical verification of digitally reconstructed radiographs (DRRs) and which can also be used to measure MTF. This is utilised to investigate the effect of CT slice thickness and pitch on MTF, and the performance of two different treatment planning systems is compared. A final illustration is included of IQWorks being used to characterise and calibrate digital display devices: a consumer grade computer monitor

and a seminar room projector. Rather more detail is presented for this last example because display devices are a critical component of any application of images in radiotherapy[240], for any imaging modality and at any stage of the imaging process, yet their full consideration in a radiotherapy context is still in its infancy.

Together, these examples demonstrate that the package can be employed to provide a consistent approach to optimising performance across the entire radiotherapy imaging chain.

5.2. Radiographic Projection Imaging

Characterisation of diagnostic radiographic systems in terms of modulation transfer function, noise power spectrum and detective quantum efficiency under standard exposure conditions is well established, being the subject of a current IEC standard[146]. State-of-the-art radiotherapy simulators are equipped with flat-panel detectors similar to those found in diagnostic imaging so the same techniques utilised with the diagnostic equipment can be applied directly to these. Although in recent years the use of conventional radiotherapy simulators has largely declined in favour of CT-simulation the same technologies are now being incorporated into linac-based image guidance systems, such as Varian's on-board imaging (OBI) and the Elekta Synergy products, which essentially incorporate a kilovoltage, diagnostic-grade imaging beam-line orthogonal to the megavoltage beam-line, with the two systems having coincident isocentres. It is therefore attractive to be able to characterise and routinely review the performance of such systems using methodologies in widespread use in the diagnostic sphere.

At the heart of the Varian OBI system is the PaxScan 4030CB detector, which was described in detail in chapter 4 where its application to cone-beam CT imaging was considered. It was noted in that chapter that the imager operated in its 'dual-gain half resolution' (DGHR) mode when acquiring CT projections. In this study the 'single-pulse half resolution' (SPHR) mode of a new OBI unit in Edinburgh will be characterised in terms of its modulation transfer function (MTF) and normalised noise power spectrum (NNPS). As described in section 4.2 this mode bins the detector elements 2×2 to give an effective matrix size of 1024×768, and is optimised to boost contrast-to-noise ratio (CNR) at the expense of spatial resolution.

In this experiment the detector was at an SSD of 175 cm and the exposures

were performed at 70 kVp, 1.2 mAs and with 0.5 mm copper filtration added to the X-ray beam. This is in-line with the methods of Dobbins *et al.[75]* and was used in preference to the IEC method because its more simplistic experimental arrangement makes it more practical to repeat on a routine basis or apply to a wider range of equipment. Furthermore, intercomparisons of the two methods have revealed that the differences in MTF and NPS yielded are insignificant[76, 78, 308].

It was found that the particular detector being considered saturated at a dose around 1.5 μGy. IEC guidance is that a detector should be characterised at 3 dose levels: its normal operating level, 3 times this level and 1/3 of this level. Because the detector saturated at around 1.5 μGy it was decided to consider 1/3 of this as a starting point, so that the study described here involved a detector dose of 0.5 μGy.

A tungsten square of thickness 1 mm and with machined edges was positioned on the surface of the detector at an angle of approximately 3° to the detector matrix. It was aligned so that one edge coincided with the central axis of the x-ray beam, as shown in 5.1, allowing an assessment there with minimal penumbra due to edge shadow from oblique rays.

In the 'EdgeLine' analysis module a Sobel edge detector with a search range of 5 pixels was used to localise the edge of the tungsten square, then the edge function sampled with a factor of 0.1, smoothed by a moving Gaussian-weighted second order polynomial and differentiated to yield the line spread function. The line spread function was then interpolated to 128 points before the Fast Fourier Transform taken to give the MTF. Finally, the MTF was rebinned to a spacing of 0.05 cycles/mm. (NB: Because the intention of this work was to characterise detector performance, with the test object positioned on the surface of the detector, results here are presented with spatial frequencies quoted in the plane of the detector rather than projected back to isocentre.)

Calculated MTFs in the X and Y matrix directions are plotted in figure 5.2, in which it is clear that there is no significant difference between the curves in either direction. The f_{50} extracted from this curve is 0.90 ± 0.05 cycles / mm. Also in the figure is a curve digitised from a similar experiment by Benitez *et al.[21]* in which the PaxScan 4030CB was also assessed. Although the measured and cited curves are comparable in shape and overall characteristics the Benitez *et al.* curve demonstrates a low-frequency drop due to veiling glare which is not observed in the measured curves, and the high-frequency performance of the Benitez *et al.* curve is slightly better than that measured. These may be

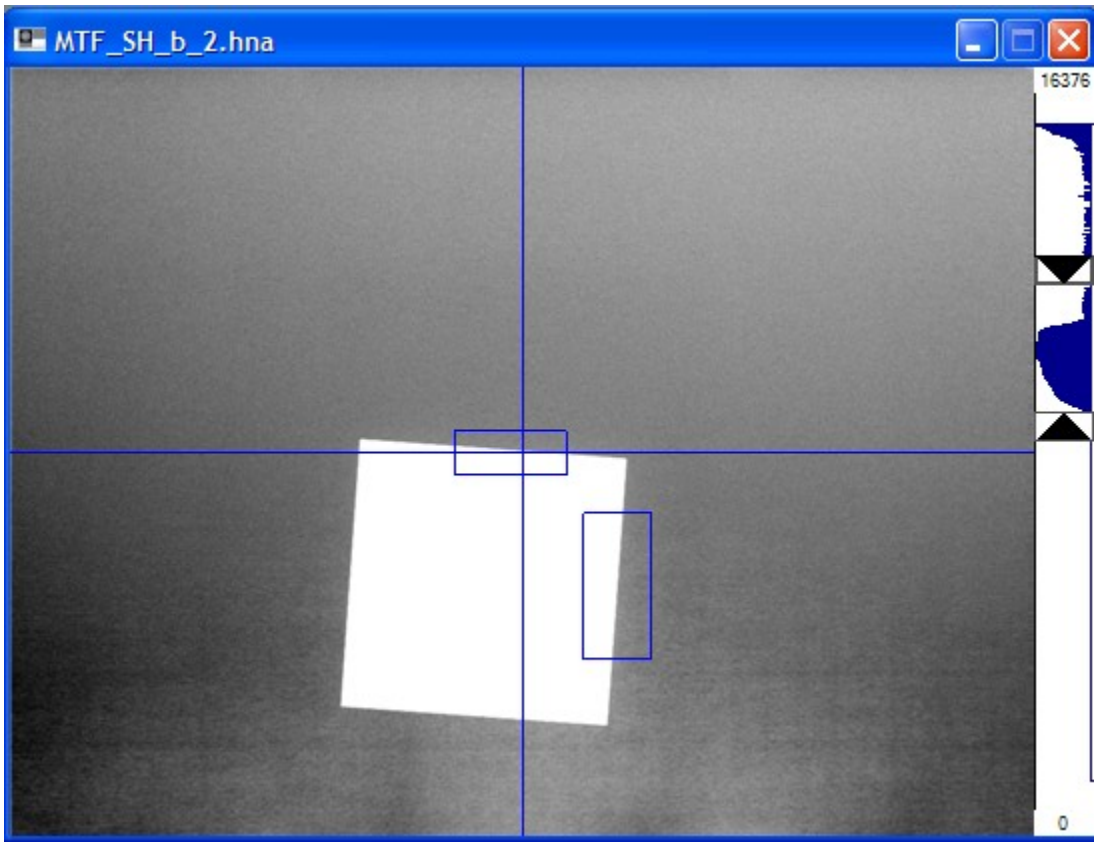


Figure 5.1. – Image demonstrating the angulation and alignment of a tungsten square when used to measure modulation transfer function (MTF). The blue cross-hairs indicate the central axis of the X-ray beam and pixels in the blue regions of interest are used to calculate the MTF. The image was acquired using the Varian PaxScan 4030CB detector of the Varian OBI unit in Edinburgh.

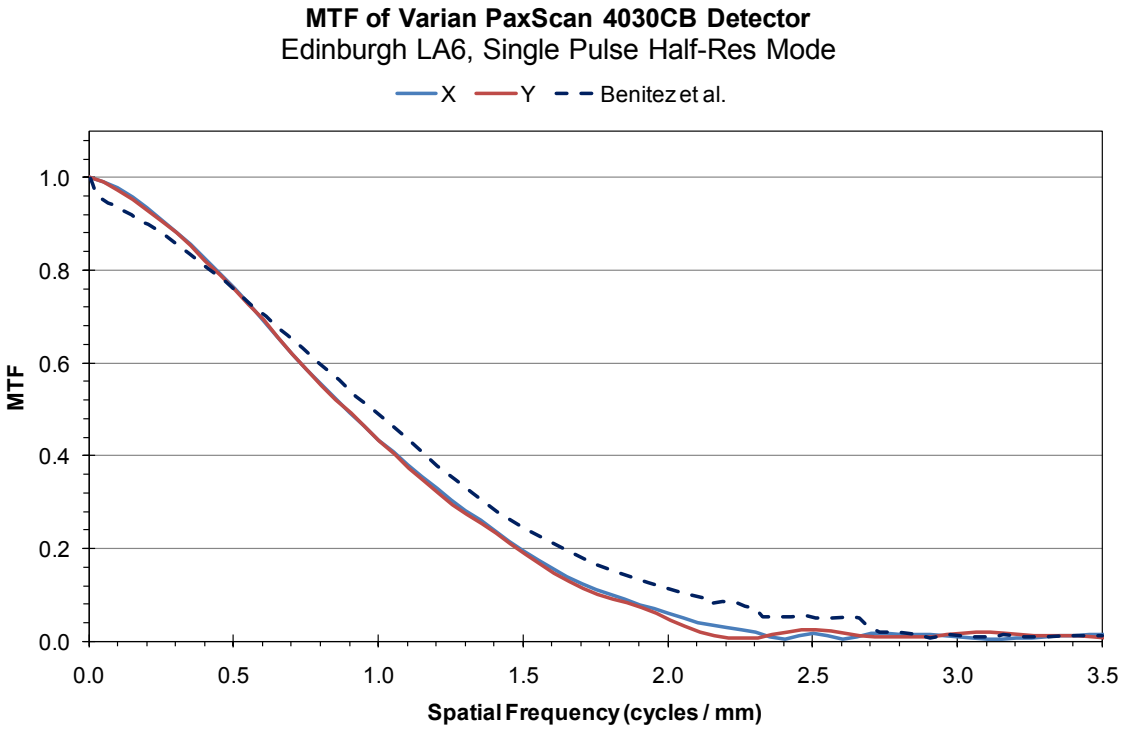


Figure 5.2. – Modulation transfer function (MTF) curves calculated in the X and Y matrix directions for the single-pulse half-resolution mode of the Varian PaxScan 4030CB detector. A curve determined in a similar experiment by Benitez *et al.[21]* is included for comparison.

explained by a different acquisition geometry (104.8 cm SSD due to mounting in a conventional CT scanner gantry, which would result in the focal spot size having a smaller effect) and the detector operating in a different mode during the Benitez team's experiment. Although not specified in their paper, because they were comparing the potential of two detectors for CBCT applications it is suspected they were characterising the dual-gain half-resolution (DGHR) mode optimised by Varian for CBCT. However, one would expect similar MTF curves for each mode because the pixel size is the same.

Normalised NPS was calculated from a flood field exposure under the same exposure conditions as described above. Regions of interest (ROIs) of 128 $\times$ 128 pixels were arranged half-overlapping across the central 95% of the field of view, resulting in 140 ROIs altogether. To minimise background trends a two-dimensional second-order polynomial fit was subtracted from each ROI and all results were normalised to the mean pixel values in the top-left ROI. The resulting X and Y direction NNPS are shown in figure 5.3, along with

curves from the work by Benitez *et al.* Whereas the Benitez team found that the NNPS was similar in each matrix direction (and therefore only present results for a single direction) the results of the current work clearly indicate higher noise power in the Y direction. Furthermore, although the NNPS curves all exhibit similar behaviour the curves measured here are approximately an order of magnitude lower than that for a lower detector dose (0.17 μGy) and are still significantly lower than that for a much higher dose (0.52 μGy). This is almost certainly due to the use of the dual-gain mode in the Benitez team's experiments – in their paper they also present results for detector doses up to ~6.4 μGy, whereas in the current work the images began to exhibit structural artefacts due to saturation at around 1.5 μGy. This is consistent with the detector appearing to have a superior dynamic range in the Benitez team's experiments, a key difference between the two acquisition modes. Furthermore, because the same dataset was used in the IQWorks validation experiment described in section B.18, in which IQWorks, QA-Distri and OBJ-IQ_reduced all yielded similar results, there is confidence that the differences observed here are not due to problems with the IQWorks algorithms themselves.

Given the significant differences in experimental conditions between Benitez *et al.* and the current work it is encouraging that the MTF and NNPS results are broadly comparable between the two and the ability of IQWorks to perform such analyses is clearly demonstrated. More work is required to fully characterise the clinical radiographic modes of the Varian OBI system in a geometrical context representative of the radiotherapy set-up verification task.

With the QEPI1 phantom being routinely used to assess the performance of megavoltage EPI it would be convenient if the same phantom could be used to test kilovoltage radiographic systems simply by rotating the gantry through 90°. In order to verify whether the QEPI1 could be used for a daily QC check of the OBI system's imaging performance and geometry, fluoroscopy images were acquired under automatic brightness control, with the phantom being aligned at isocentre exactly as for megavoltage imaging, as described in section 3.5 and the detector at the default source-detector-distance (SDD) of 150 cm. Images were analysed using a similar analysis tree for megavoltage imaging, and the resultant average MTF curve is plotted in figure 5.4, where the error bars indicate ±1 SD in the MTF values measured across the field of view. The curve from a 1 mm tungsten square aligned at isocentre has also been plotted for comparison because the tungsten square is the generally accepted test object for fundamental radiographic spatial resolution assessment. From the figure

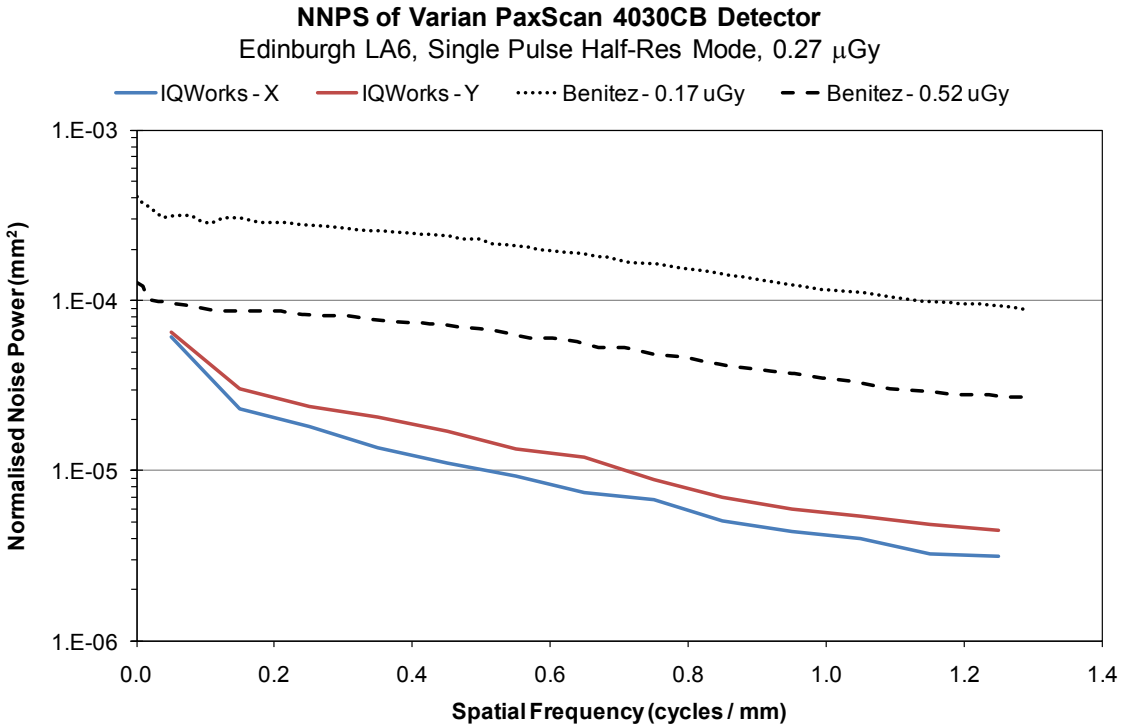


Figure 5.3. – Normalised noise power spectrum (NNPS) curves calculated in the X and Y matrix directions for the single-pulse half-resolution mode of the Varian PaxScan 4030CB detector. Curves determined for similar dose exposures in an experiment by Benitez *et al.[21]* are included for comparison.

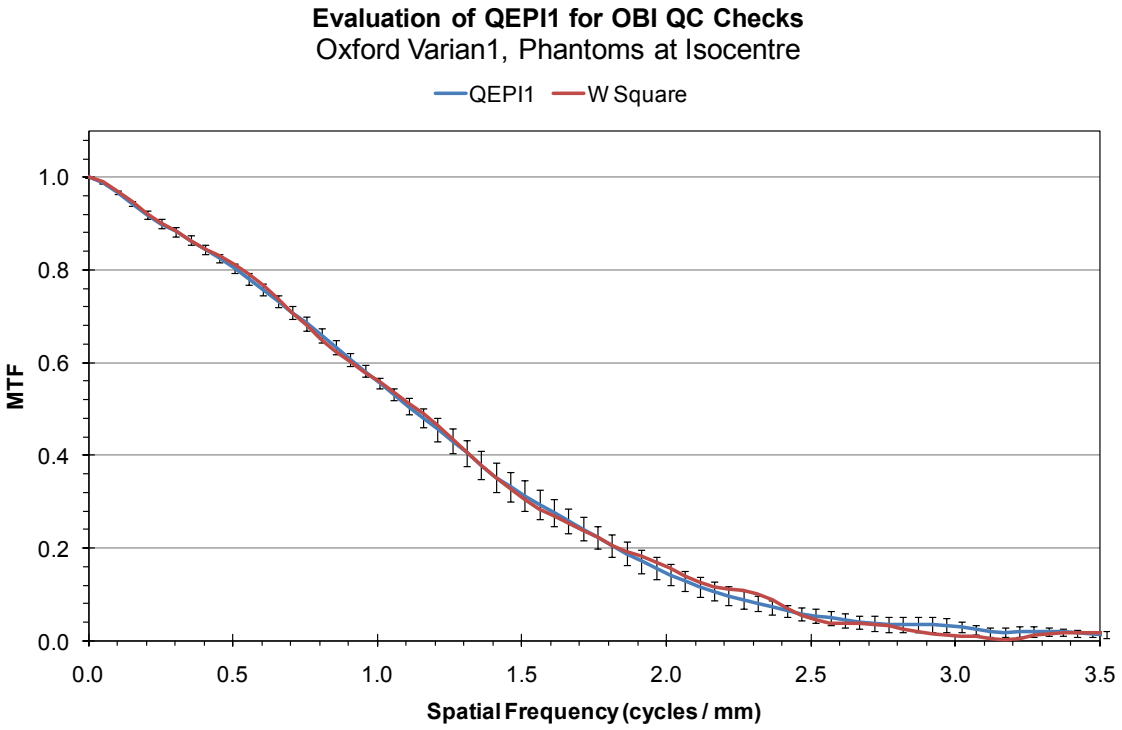


Figure 5.4. – Modulation transfer function (MTF) curve derived from a Varian OBI fluoroscopy image of the QEPI1 phantom aligned at isocentre. Error bars indicate ±1 SD about the mean MTF curve calculated across the field of view. The curve obtained from a tungsten square imaged using the same geometry is included for comparison.

it is clear that both curves agree within the limits of experimental uncertainty, thus suggesting the QEPI1 phantom can reliably be used for routine QC checks of the radiographic components of the OBI system.

Taking both curves together, the f_{50} was measured as 1.60 ± 0.05 cycles/mm. This is consistent with the curve above using the tungsten square on the surface of the detector. In the previous experiment the detector SDD was 175 cm. Correcting the current measurement for geometric magnification suggests that the measurement on the surface of the detector would yield an f_{50} of 1.0 cycles/mm. This is close to the 0.90 cycles / mm measured above and it is encouraging that consistent results can be obtained using either phantom.

5.3. Digitally Reconstructed Radiographs (DRRs)

A new phantom[290] was developed as part of this work to harness the potential of the 'DRR Geometry Checker' described in section B.20. The phantom consists of high-contrast ball-bearings suspended in a tissue-equivalent base material and is flexible in that the number, locations, diameters and materials of the ball-bearings, as well as the choice of base material, can be varied depending upon the intended application. For example, the phantoms shown in figure 5.5 contain a combination of 1 mm diameter aluminium ball-bearings and 0.4 mm diameter tungsten ball-bearings embedded in either wax or a polyester resin. Whereas the wax is easy to use to build prototype phantoms containing different types of ball-bearings the polyester resin is more durable and can be machined, so is more appropriate for a permanent phantom intended for use in a routine QA programme. However, the resin phantom is considerably more difficult to manufacture.

Coordinates of ball-bearings in a DRR image can be predicted for any combination of gantry angle, couch rotation and isocentre shift using the 'DRR Geometry Checker' module, with the predictions then verified through observing that the cross-hairs overlaid on the image indeed match the locations of the balls. In addition, the 'Impulse MTF' analysis module described in section B.17.4 can be used to calculate the MTF from a bearing's point spread function, allowing the influence of CT acquisition settings on DRR spatial resolution to be investigated.

When constructing the DRR phantom it was found that even small imperfections near the ball-bearings could influence the CT scans and thus the resultant DRRs. Boring holes through which the balls were inserted was therefore not a suitable construction technique because it was almost impossible to avoid residual pockets of air when backfilling the bore holes with base material. Instead, rather than precisely positioning the balls during manufacture, it was determined that a more reproducible finish could be achieved through building the phantoms in layers and roughly placing the balls on the surface of a still soft layer before the next layer was poured. Once the phantom had hardened, and the edges squared off through machining, the precise locations of the balls could be identified through thin-slice CT scanning.

Another challenge was striking a balance between the contrast and pinpoint localisation of a ball-bearing and its size and the photon starvation or beam hardening artefacts caused by the dense material. Through experimentation it was discovered that the relatively large 1 mm diameter aluminium bearings

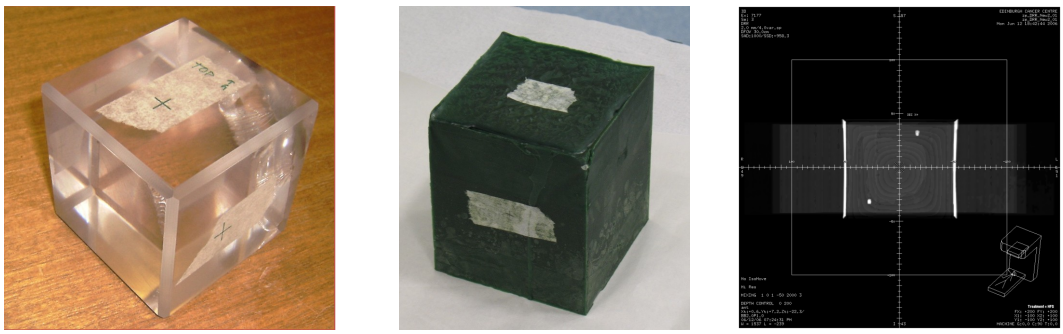


Figure 5.5. – Photographs of polyester resin (left) and wax (middle) prototypes of the new phantom for checking Digitally Reconstructed Radiograph (DRR) geometry. A DRR generated in Advantage Sim, on which the projections of two ball-bearings are clearly visible, is on the right hand side.

could be used for geometrical evaluations under all acquisition conditions but the very small tungsten bearings were more appropriate for detailed MTF analysis, although these could be difficult to visualise without optimising the DRR parameters.

Although it is recommended that regular quality assurance is undertaken of the DRR generation process, including data transfer mechanisms[154, 180], and furthermore that DRRs are optimised as part of the overall set-up verification strategy[284], once a DRR platform is introduced into clinical use it is generally accepted that it 'just works'. However there are real clinical scenarios where DRRs can be non-optimal but where the limitations may not immediately be apparent to the complex content of patient images and the way in which treatment planning or virtual simulation systems handle images. For example, if there is a glitch in the network transfer of CT slices between the CT scanner and treatment planning system then the volumetric dataset utilised for treatment planning and DRR generation may have slices missing. Although to a human observer a gap in a stack of slices may be obvious, treatment planning systems tend to deal with this without warning the operator because it could be a deliberate feature of the particular scanning protocol (along with multiple groups of different slice thickness, non contiguous slices, etc.). One strategy for dealing with missing data is to generate a 'virtual' slice at the gap location which is interpolated from the real slices either side. If undetected this can have enormous consequences if the slice is at the very edge of the CTV or PTV and thus relied on for outlining or field placement decisions. Because structures inside the body vary smoothly with position interpolated data can appear very realistic, whereas in a geometrical phantom with sharp edges interpolations

tend to be less successful.

A screenshot demonstrating IQWorks being used to analyse an image of one of the DRR phantom prototypes was shown in figure B.17. Further examples of IQWorks being used to predict the locations of the ball-bearings in DRRs generated by Varian Eclipse and GE Advantage Sim over a range of geometrical scenarios are included in figure 5.6. One considerable advantage over other systems, such as the Modus Quasar MLC Beam Geometry Phantom[98, 253], is that a whole range of scenarios can be quickly investigated from just a single CT scan of the phantom, or at worst from a single placement of the phantom and multiple CT scans acquired one after the other. Other phantoms require time-consuming realignment of the phantom prior to each CT scan, limiting their practical implementation into a routine QA programme. IQWorks and the new DRR phantom therefore provide a powerful and flexible platform for verifying DRR geometrical accuracy under different CT acquisition and DRR generation conditions.

When this work was being undertaken two virtual simulation systems were available in the Edinburgh Cancer Centre, Varian Eclipse and GE Advantage Sim. Each system allowed the user to vary how DRRs were generated through modifying a number of parameters. It was desirable to compare the fundamental performance of DRRs from the two systems as well as better understand the influence of CT acquisition settings on DRR spatial resolution. One of the new DRR phantom prototypes was scanned multiple times using the standard Edinburgh radiotherapy head and neck protocol. This was a helical technique with settings of 120 kVp, 100 mA, 0.8 s rotation and with data being acquired and reconstructed over a 35 cm field of view. Between scans the CT slice thickness was varied in steps from 1 mm to 10 mm, and whole sets of scans were acquired at pitches 1.0 and 1.5. Scanning at a pitch of 1.5 reduces the overall time required for the scan by approximately one third, thus making it more practical to acquire thinner slices, but there was concern that this would also result in smearing of small objects and therefore a corresponding loss of spatial resolution in the longitudinal direction in DRRs. For each set of slices a DRR for an anterior beam was generated in each system, with depth control options being applied to minimise the loss of contrast due to material over or underlying one of the 0.4 mm tungsten bearings. The DRRs were then exported as DICOM files, imported into IQWorks and the 'PSF Analysis' module employed to calculate the MTF in the X and Y matrix directions in the DRR. These correspond to the left-right (i.e. transverse) and inferior-superior (i.e. parallel to the

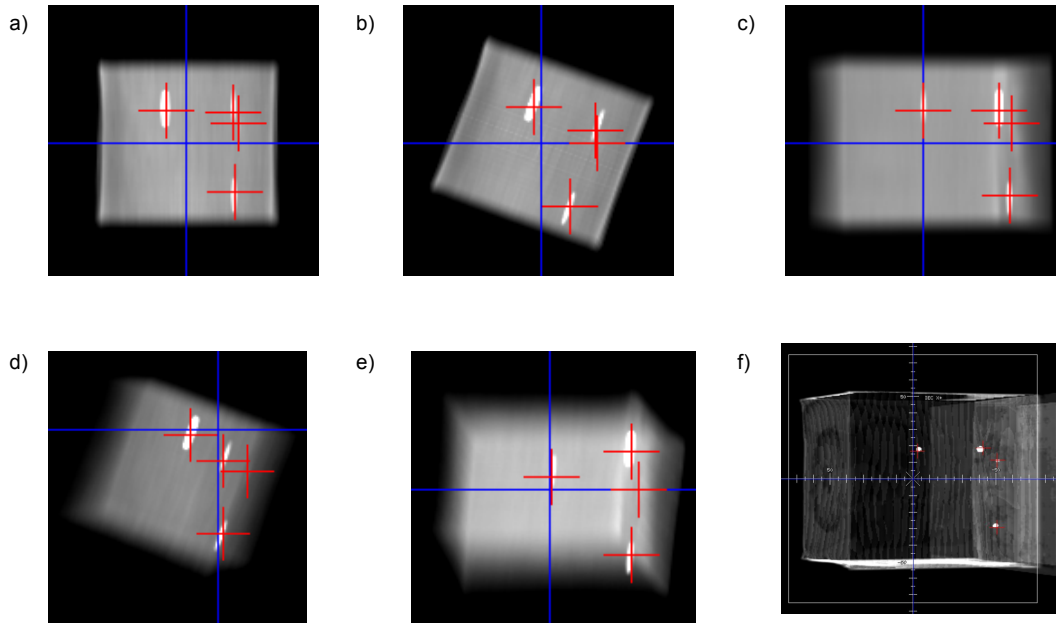


Image	System	Gantry Angle (°)	Couch Rtn (°)	Isocentre		
				X (mm)	Y (mm)	Z (mm)
a	Eclipse	0	0	0	0	0
b	Eclipse	0	340	0	0	0
c	Eclipse	75	0	0	0	0
d	Eclipse	75	340	0	0	0
e	Eclipse	15	340	10	20	30
f	Adv Sim	75	0	0	0	0

Figure 5.6. – Examples of IQWorks and the DRR phantom being used to verify DRRs generated by Varian Eclipse and GE Advantage Sim for the beam geometries indicated in the table.

couch motion) directions in a patient.

Plots of MTF f_{50} for the two systems are presented in figure 5.7, where each system is considered separately and the effect of the two pitch settings is compared, and figure 5.8 where the systems are directly compared against each other on the same plots. Error bars are ±1 SD of the average of three repeated measurements. As expected, for both systems and pitch settings, as CT slice thickness decreases, f_{50} in the Y direction improves. However, also for both systems and pitch settings, there is at the same time a corresponding deterioration in f_{50} in the other direction. This is an important result because it demonstrates that DRR optimisation is not as straightforward as opting for the thinnest possible CT slices, with instead a balance between spatial resolution in each direction probably being more appropriate. The experiment here suggests that a slice thickness of between 1 and 2 mm may be optimal.

Also as expected, and for both systems, the higher pitch setting results in a loss of resolution in the Y direction. However, two other interesting results are that resolution in the X direction is also degraded (which may be surprising because this is normal to the scan direction, and so less influenced by pitch) and that the Y direction degradation is worst for the thinnest slices. In fact, from these results, the boost in f_{50} from using a 1 mm slice rather than a 2 mm one is lost if the pitch is increased from 1.0 to 1.5, perhaps to make the thin slice acquisition viable! Careful consideration must therefore be given to CT scan acquisition parameters when aiming to generate DRRs for treatments requiring tight set-up correction protocols.

Overall, at a pitch of 1.0 the f_{50} performance of Eclipse and Advantage Sim was comparable, with Eclipse performing marginally better in the X direction. However, at a pitch of 1.5, although the two systems perform similarly in the Y direction, Eclipse is considerably better in the X with the discrepancy being worse for thinner slices. It is uncertain why this is the case but it is likely a result of the different algorithms employed by each system to generate their DRRs. This is partly illustrated in figure 5.9, which shows the 2D point spread function of the ball bearing for 1 and 10 mm slices in each system. Whereas Eclipse has a relatively smooth PSF in both directions, the Advantage Sim DRRs contain significantly more structure. There is also evidence of saturation-like truncation of the Advantage Sim PSF for the 10 mm slice. Finally, the Advantage Sim DRRs are generated over a finer grid, set by the resolution of the display device when the DRRs are calculated, whereas Eclipse DRRs are always over a 1024×1024 matrix over a specified FOV. Advantage Sim DRRs are therefore susceptible

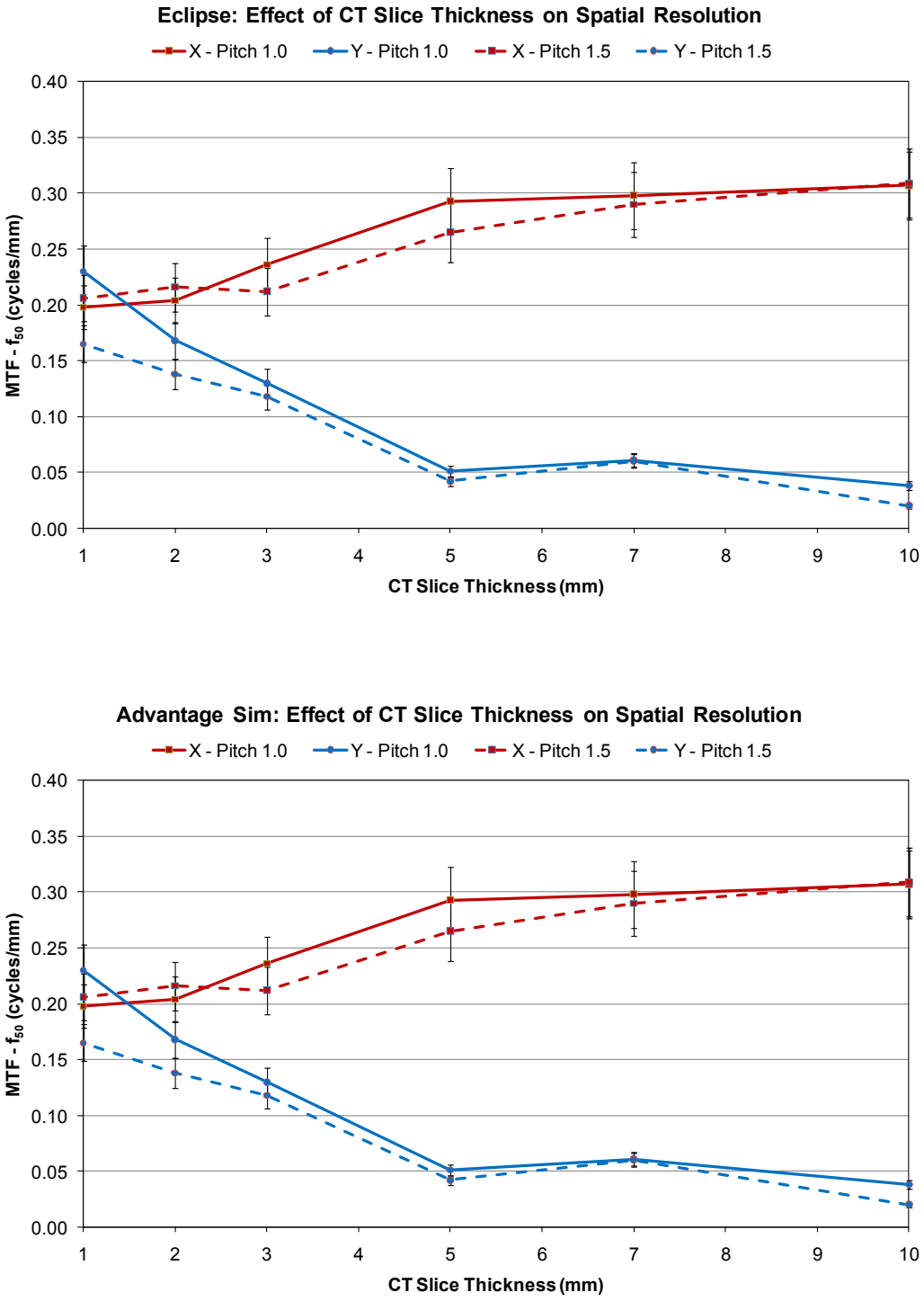


Figure 5.7. – Effect of CT slice thickness and pitch on the modulation transfer function (MTF) f_{50} of DRRs generated by Eclipse (top) and Advantage Sim (bottom).

to local noise variations (as evidenced by the rough edges in the images in the figure) which may in turn influence the MTF calculations.

Although the new DRR phantom is optimised for this application, other phantoms can also be utilised with IQWorks to assess DRR spatial resolution. Analyses have also been performed on DRRs generated of one of the Catphan 504 MTF beads, as illustrated in figure 5.10. For a new clinical trial in Oxford there was an aspiration to acquire the planning CT scan under breath hold, but there was concern that patients would be unable to hold their breath for the duration of the existing 2.5 mm, pitch 1.0 scanning protocol. It was suggested that instead the patients should be scanned using 3.75 mm slices, with all other parameters remaining identical, but that the images sent to the planning system would be reconstructed at a spacing of 2.5 mm. IQWorks and the Catphan MTF bead were used to compare the proposed and existing protocols in terms of the full-width at half-maximum (FWHM) of the DRR PSFs in the Y direction. The original protocol had a FWHM of 2.8 ± 0.3 mm and the proposed protocol one of 4.0 ± 0.7 mm. Based on the results of this study it was felt that the new protocol did not yield DRRs of sufficiently high resolution to match the 3 mm action level of the set-up correction protocol intended for use during treatment. It was therefore decided to continue using the existing scanning protocol.

With high resolution planar or volumetric imaging becoming readily available at the treatment machine for patient set-up verification it is important the DRRs or CT scans against which these are being matched are optimised so that they are not the limiting step. However, whereas there is a drive generally to improve the precision of set-up verification there is often little effort invested in reviewing pre-treatment CT protocols or DRR generation settings alongside this. Together, IQWorks and the new DRR phantom described above provide a quick and robust mechanism for objectively characterising DRR performance and addressing real clinical optimisation questions[284].

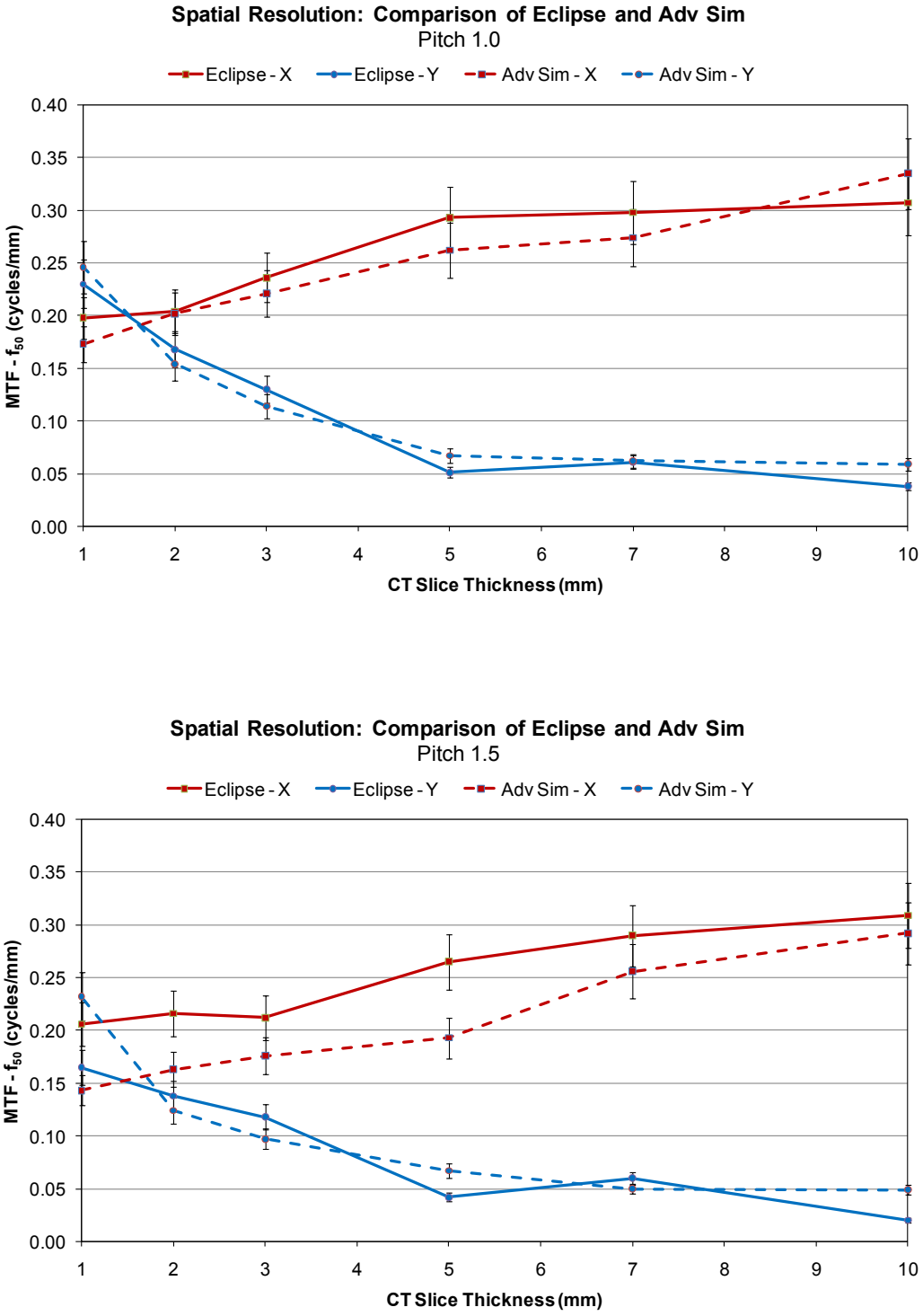


Figure 5.8. – Comparison of Eclipse and Advantage Sim MTF f_{50} performance for CT scans with a pitch of 1.0 (top) and 1.5 (bottom).

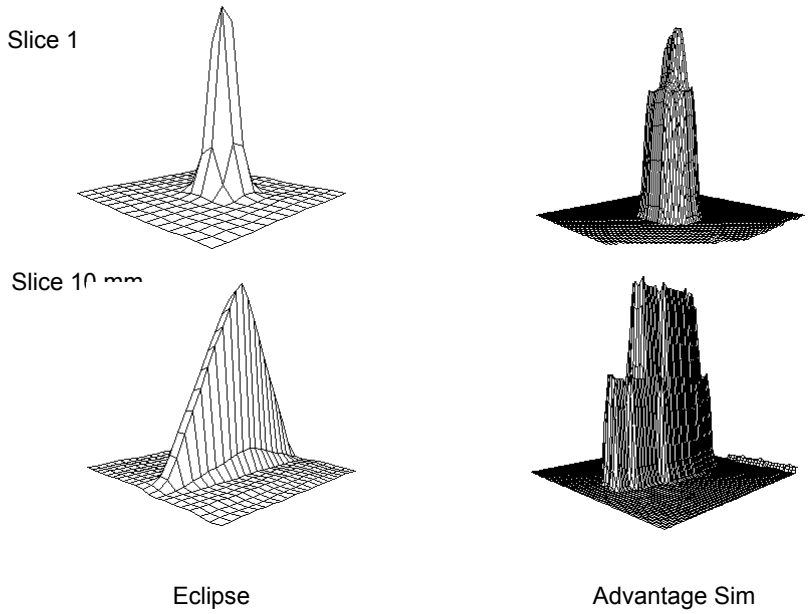


Figure 5.9. – Comparison of two-dimensional point spread functions for DRRs generated from thick (10 mm) and thin (1 mm) CT slices by Eclipse and Advantage Sim.

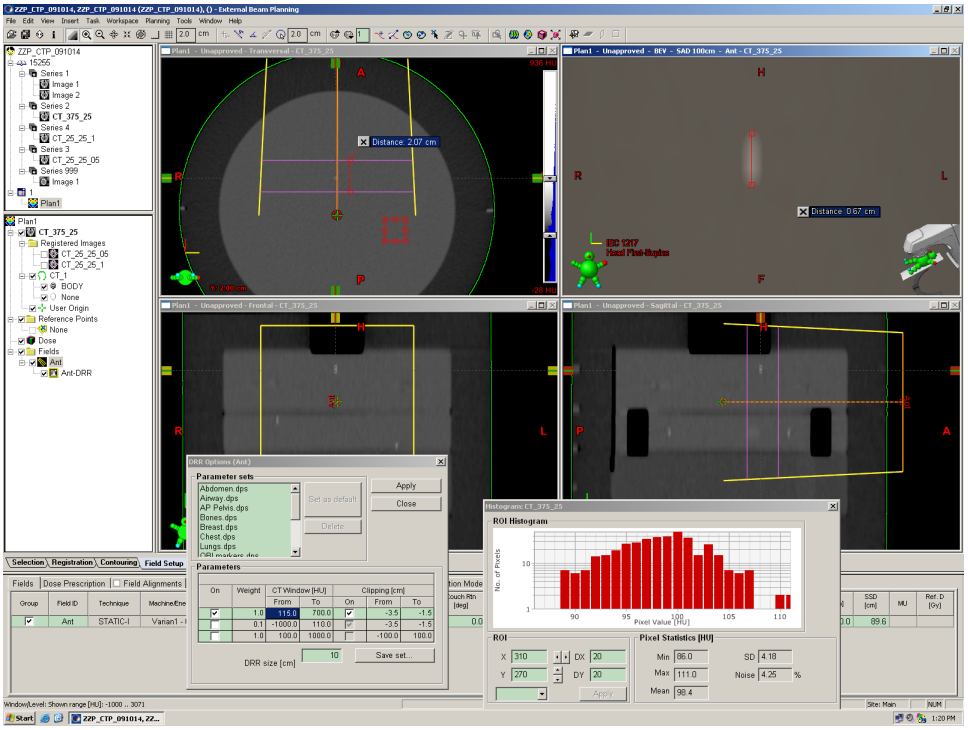


Figure 5.10. – Screenshot from Varian Eclipse in which the DRRs are generated from a volume localised around one of the impulse beads.

5.4. Display Devices

Across the entire radiotherapy patient pathway clinical decisions are increasingly being made directly from digital images displayed on a computer monitor or by a projector in a seminar room. With the drive towards paperless or paper-light clinical environments, and the widespread acceptance of digital signatures, it is now not uncommon for a radiotherapy treatment plan to be prepared, sanctioned, delivered and verified without hard copies ever being made of the images which were relied upon at the different stages of the process. At the very end of the image presentation chain, display devices can potentially have an enormous influence on how a human observer interprets an image and therefore the clinical decision which is arrived at. Although quality assurance and optimisation of display devices is recommended both by equipment vendors[380] and in national guidelines[240, 305] even procedures for basic testing of displays will be unfamiliar to many radiotherapy departments.

Minimum specification of displays for radiotherapy applications are generally less than for diagnostic ones[375]. However, it is still important that all displays in a radiotherapy department are part of a quality assurance programme. It is the author's experience that perceived modality problems – or even the choice of portal imager used to acquire images for a particular group of patients – may actually be the result of the configuration of the display device on which the images are regularly reviewed. The ARIA radiotherapy network in the new Oxford Cancer Centre is typical of that in many contemporary centres and enables imaging tasks for treatment planning and set-up verification to be performed on any number of equivalent workstations across the entire geographical department. Perhaps more important than absolute performance – as long as a particular display device is 'good enough' for the radiotherapy application – is that all devices which may be used for a particular application present the images to the user in a consistent fashion, so that the same clinical decision would be reached regardless of where in the department the assessment was performed. As discussed in section B.21 it may also be necessary to take into consideration ambient lighting conditions.

Some examples are presented below of how display assessment methodologies from the diagnostic imaging arena may be applied in a radiotherapy environment. A discussion of the numerical implementation of these in IQWorks is included in Appendix B.21.

Given that display QA is a relatively new concept for radiotherapy it is important to adopt a pragmatic approach in line with the requirements of the

clinical application. Ensuring LCD displays operate at their native resolution, and that the contrast and brightness are set consistently, can easily be achieved using standard test patterns[106, 305, 382]. However, verifying consistency of characteristic curves requires access to measurement instruments and this may be difficult due to cost implications.

Photometers generally operate in two modes: telescopic, where the instrument is pointed at a target from a distance, and in contact, where the instrument is placed on the surface of the display device. Some devices are configurable to operate in either mode. Telescopic photometers tend to be very expensive (thousands of pounds) but have the advantage that their readings take into consideration ambient light reflected from the target. However, through careful placement of the display device, and information gained from an environmental light survey, it is also possible to account for ambient light even when using a contact photometer. Nevertheless, medical or scientific grade photometers still tend to be prohibitively expensive for use in an emerging application such as radiotherapy.

One of the accepted 'gold standard' telescopic photometers is the IBA LXPlus (now superseded by the LXCan) and the accepted standard contact photometer is the Image Smiths' VeriLUM. Both of these are shown in figure 5.11. However, through investigating these instruments it was discovered that the VeriLUM is based on an X-Rite chromaticity sensor, and that X-Rite also manufacture considerably less expensive devices based on similar technology for the professional graphic design market. One such device is the X-Rite eye-one display 2, which is shown in figure 5.11d. A version of this instrument with minimal bundled software retails for just over £100. An experiment was performed to determine whether the eye-one display 2 would be suitable for radiotherapy display device assessment.

At the time of the experiment, all new Varian imaging workstations were delivered with Dell 20" LCD displays. One of these was chosen as a reference device and the IQWorks 'Display Assessment' module used to measure the characteristic curve, first using the LXPlus 'gold standard' then with a loan eye-one display 2. A comparison was also made on a typical Belinea 21" CRT display and the results of both experiments are presented in figure 5.12. It is immediately clear that there is excellent agreement between the two instruments when used to assess either display device, with the curves overlying so closely as to be indistinguishable from each other.

Both the eye-one display 2 and LXPlus also have illuminance sensors which

Figure 5.11. – Photographs of different photometers. a) IBA LXPlus in contact mode; b) IBA LXPlus in telescopic mode; c) Image-Smiths' VeriLUM contact pod; d) X-Rite eye-one Display2 contact pod.

are useful for projector characterisation and ambient light surveys. Another intercomparison was performed but this time using a projector as a reference light source and with the detectors in the geometry shown in figure 5.13. From the plot in the figure it is clear that the illuminance response of the eye-one display 2 is also very good and it is likely this could be improved through the application of a simple calibration factor.

There are sometimes discussions regarding how long a display device should be allowed to warm up before being used for a clinical application, or whether the colour performance of the display is important. Through repeated automated runs the eye-one display 2 and IQWorks can be applied to the first question, and an addition to the analysis module which allows extraction of chromaticity information can address the second. These are active areas of investigation and the plots in figure 5.14 illustrate the detailed information which can be obtained.

It is clear from these investigations that the eye-one display 2 is an inexpensive yet powerful tool which has many potential applications as part of a radiotherapy imaging QA programme.

One method for monitoring the consistency of display device performance is to compare the response of any given device to the DICOM Grayscale Stand-

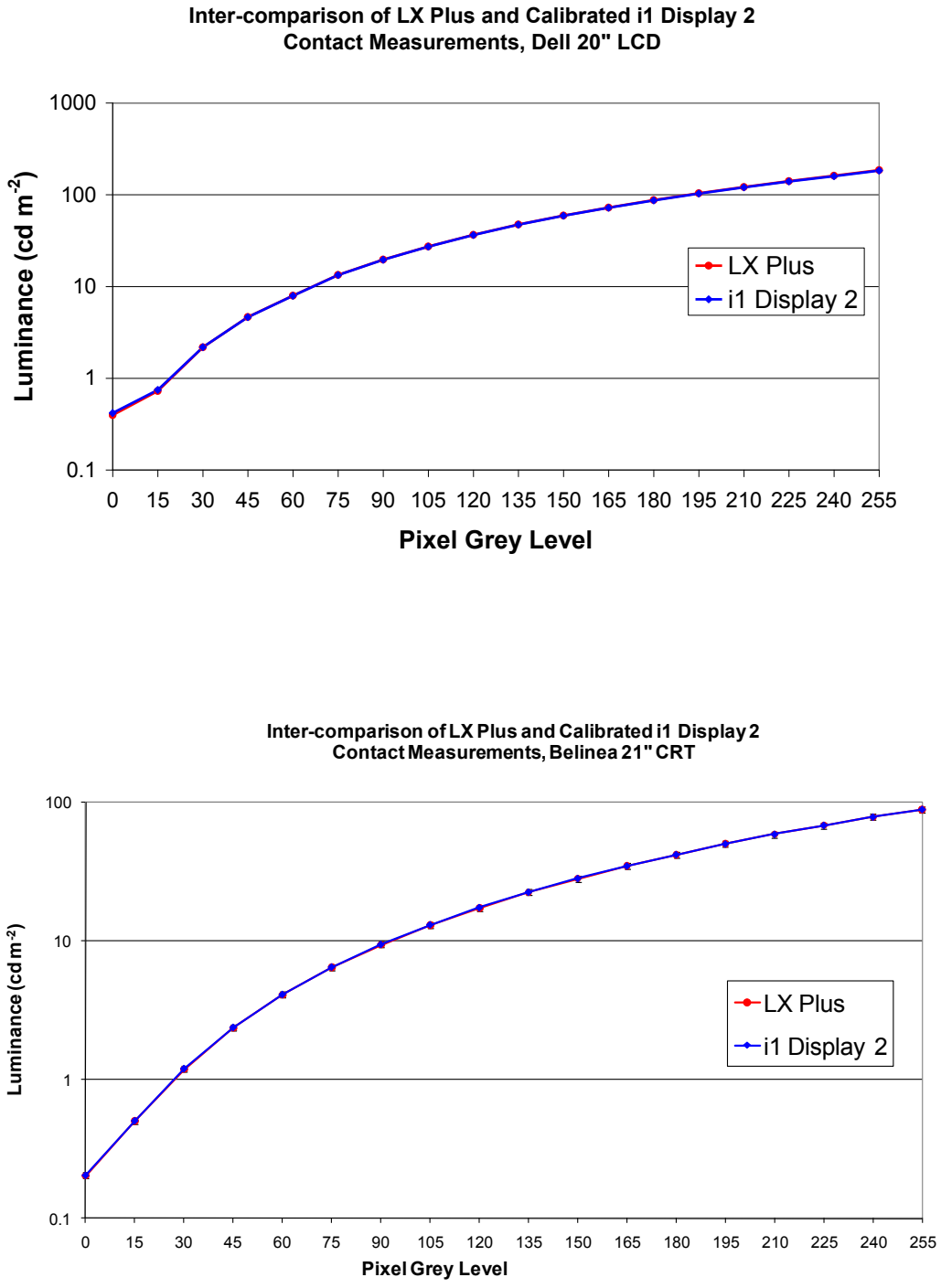


Figure 5.12. – Results of using IQWorks to measure the characteristic curve of two displays using the IBA LXPlus and X-Rite eye-one Display 2 detectors. Top: LCD display. Bottom: CRT Display.

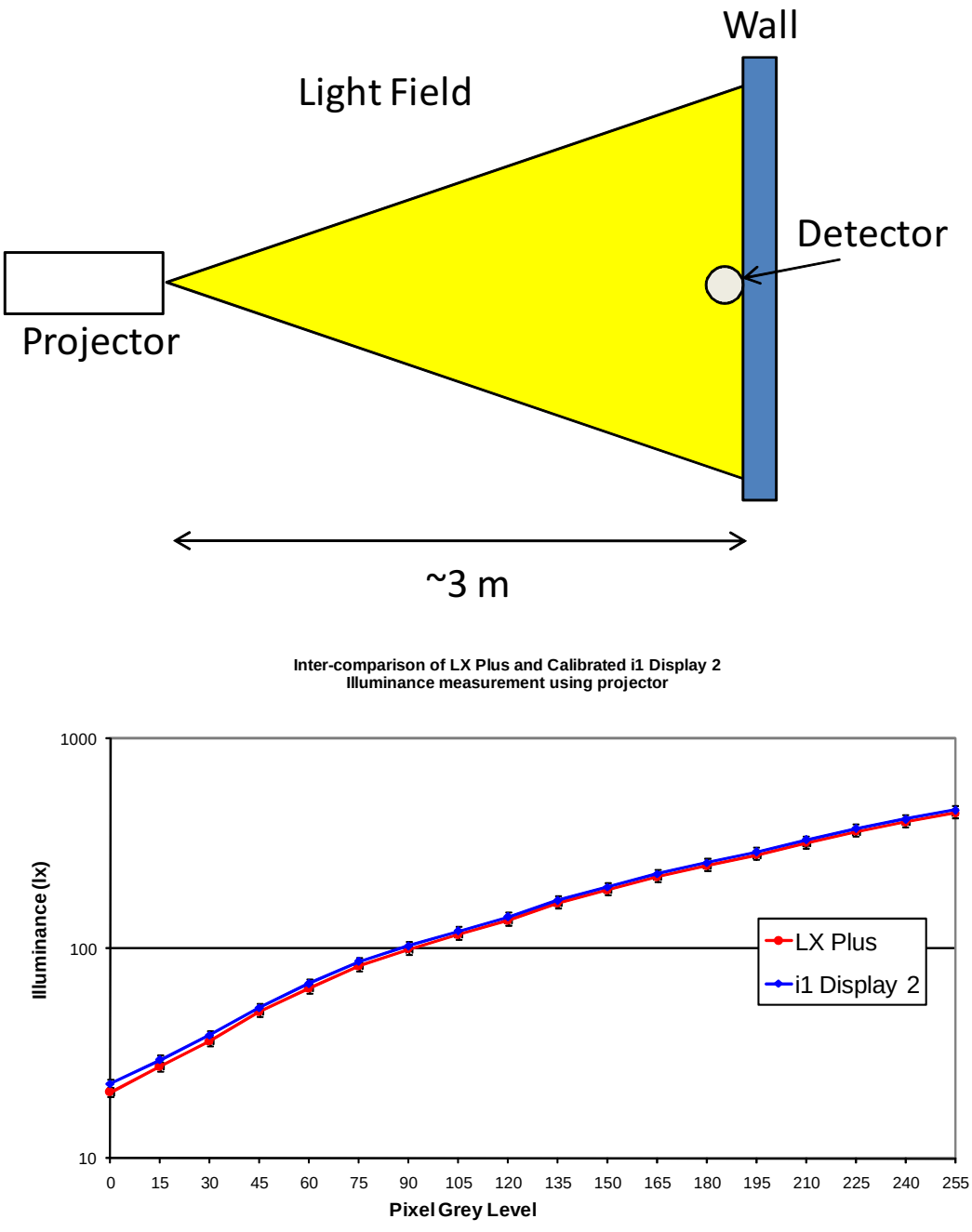


Figure 5.13. – Comparison of LXPlus and i1Display2 for illuminance measurements. Top–Measurement geometry. Bottom–Results of comparison.

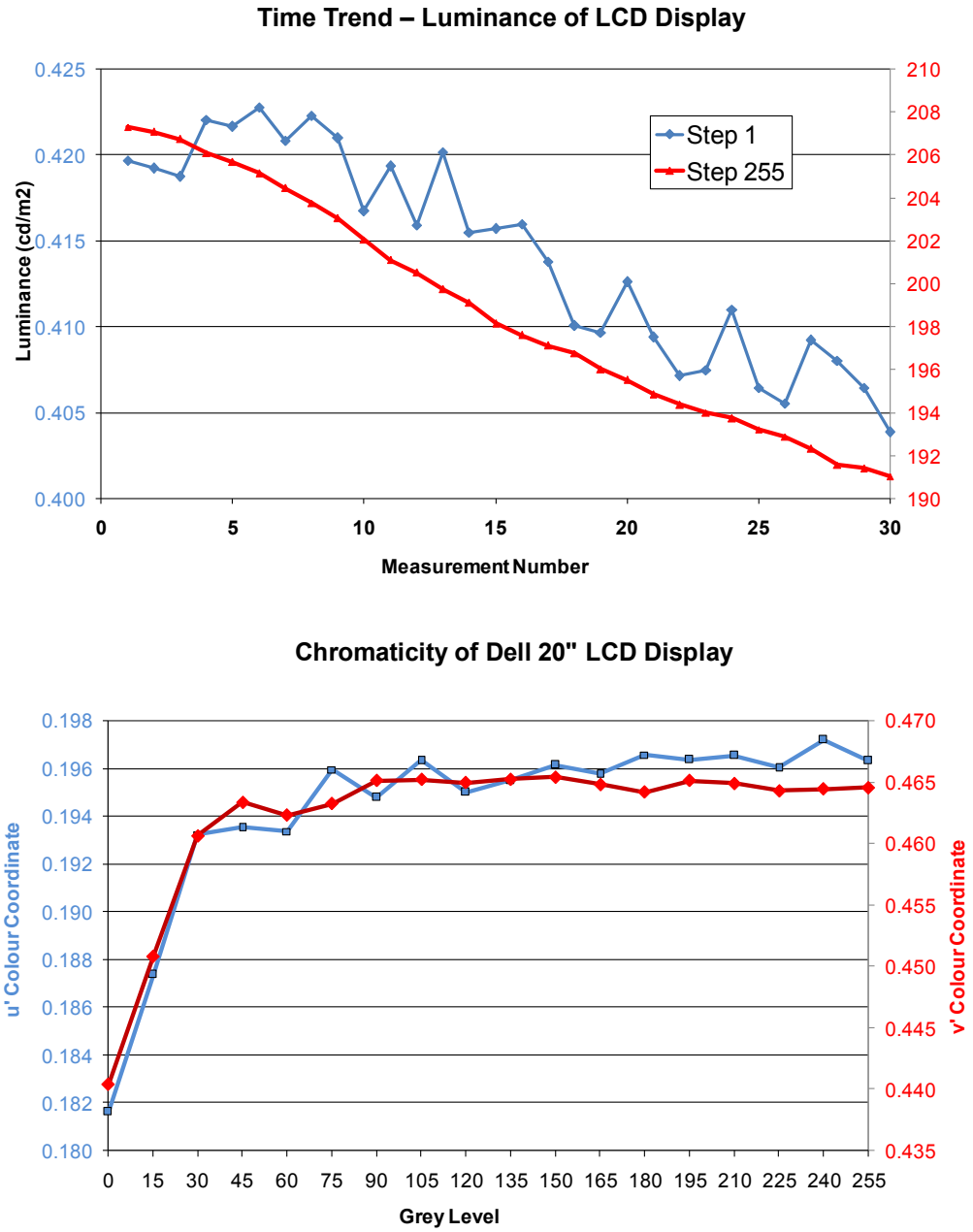


Figure 5.14. – Other display assessments using IQWorks and the eye-one Display2 detector, both of a Dell LCD display. Top–Warm-up trends of two luminance steps in TG18 test pattern. Bottom–Chromaticity characteristics

ard Display Function (GSDF)[374]. . In another experiment, the LXPlus was used to measure the characteristic curve of the Dell display and calculate the look-up table (LUT) necessary to calibrate it against the GSDF. The assessment was then repeated with the LUT activated to verify the accuracy of the calibration. Altogether, the calibration and verification process took less than five minutes. In figure 5.15 the characteristic curve before and after calibration is compared against the GSDF using the methodologies in the AAPM Task Group 18 Report[305]. In the top graph in the figure luminance is compared directly, whilst in the lower graph the first differentials of the curves are compared (known as the "JND Contrast Curves"). It is clear from the results that not only is the IQWorks assessment process viable, but the resultant calibration brings the display almost completely into the GSDF ± 20% tolerance band for a 'Class 2' display device. This is defined in the TG 18 Report as being suitable for radiology review applications[305].

It was also attempted to calibrate a seminar room projector against the GSDF. The results pre and post calibration in terms of output luminance per just-noticeable difference (JND) index and the JND contrast curves are shown in figure 5.16. Again, this exercise was successful, with the projector falling completely within 'Class 2' device tolerances. These results are encouraging because they demonstrate that, with relatively little effort, it is possible to standardise radiotherapy display device performance across disparate devices, in-line with diagnostic imaging practice and national guidance[240].

Finally, the eye-one display 2 was employed to undertake surveys of environmental light levels in the two treatment planning rooms in Oxford. An unobtrusive monitoring station was set-up alongside a planning workstation, as shown in figure 5.17, and light levels monitored over the course of a day. On a second day the exercise was repeated next to an image analysis workstation in the other room. From the results shown in figure 5.18, although fluctuations are apparent in the light levels in either room, the levels in the 'Open Lab' are almost an order of magnitude higher than in the 'Planning Room'. Work is clearly required to better standardise radiotherapy image viewing conditions. However, the important result is that using IQWorks and inexpensive detectors such as the eye-one display 2 it is definitely possible to actively monitor and manage the ambient lighting for radiotherapy imaging applications.

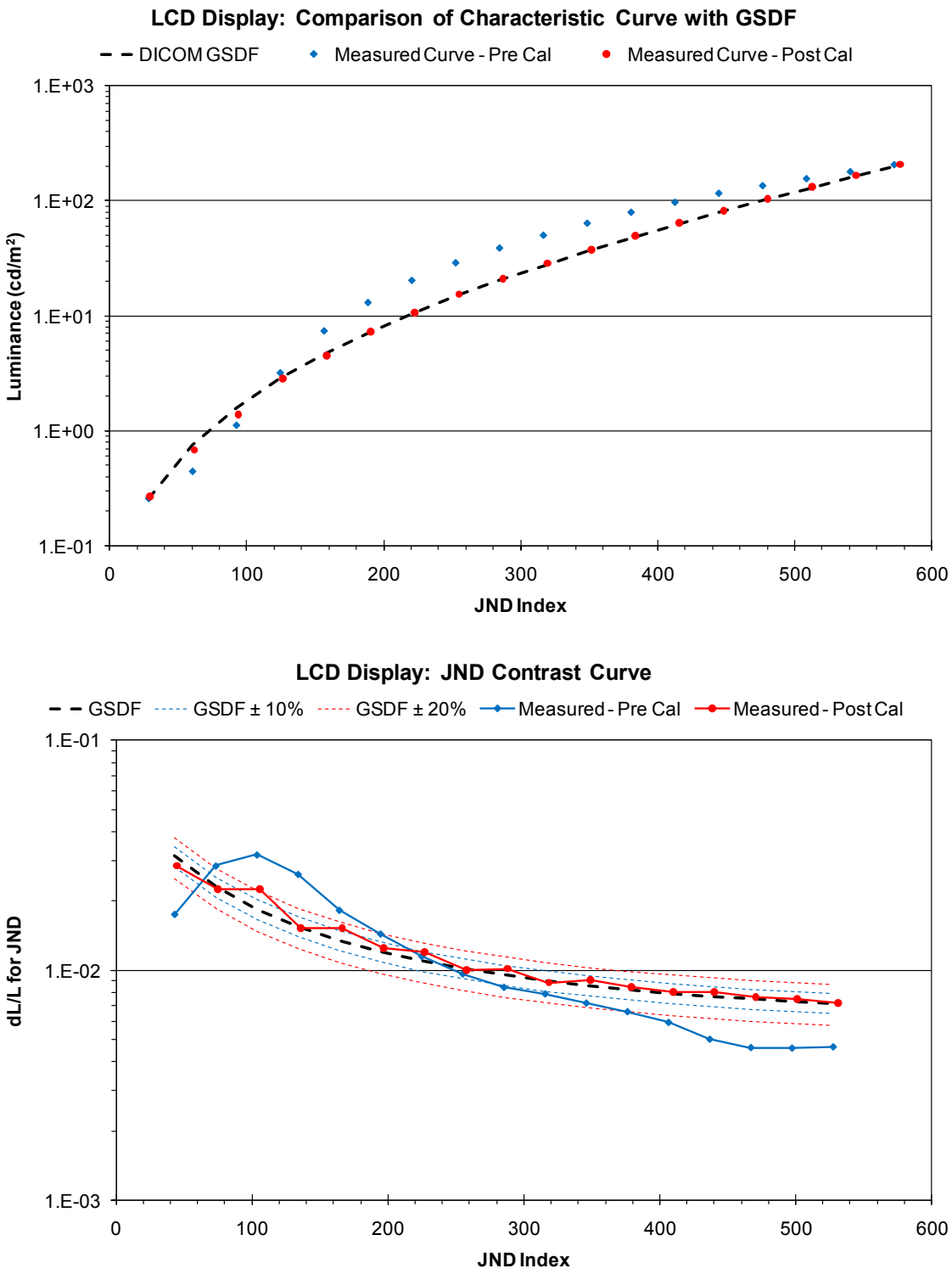


Figure 5.15. – LCD Display Calibration. Comparison of display characteristic curve against DICOM Greyscale Standard Display Function (GSDF) before and after calibration by IQWorks. Top: Luminance Curves. Bottom: Just-Noticeable Difference (JND) contrast curves. Measurement error bars are about the size of the data points.

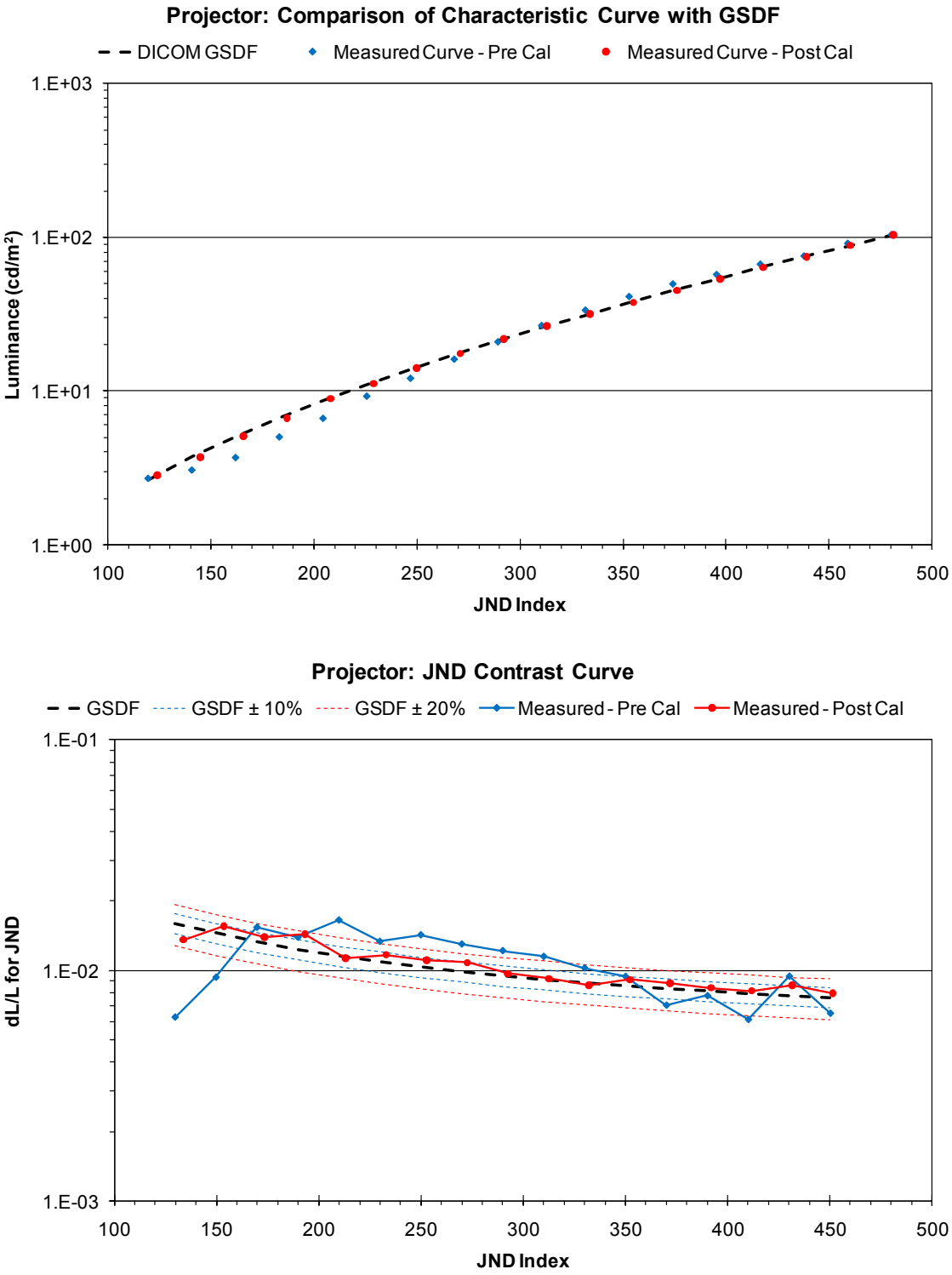


Figure 5.16. – Projector Calibration. Comparison of display characteristic curve against DICOM Greyscale Standard Display Function (GSDF) before and after calibration by IQWorks. Top: Luminance Curves. Bottom: Just-Noticeable Difference (JND) contrast curves. Measurement error bars are about the size of the data points.

Figure 5.17. – IQWorks and the eye-one Display2 photometer being used to datalog environmental light levels in the Oxford radiotherapy treatment planning room.

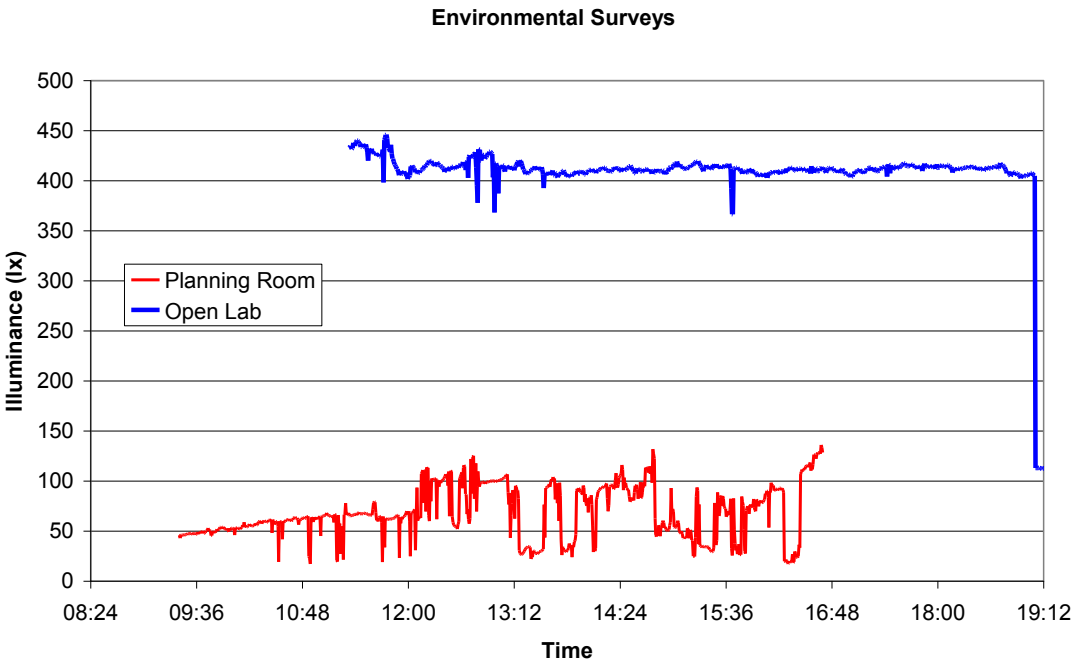


Figure 5.18. – Comparison of ambient light levels in two treatment planning areas in Oxford Cancer Centre.

5.5. Conclusion

Throughout this chapter the ability of IQWorks to analyse images of test phantoms for a range of imaging modalities has clearly been demonstrated.

A preliminary characterisation of the PaxScan 4030CB kilovoltage detector associated with the Varian OBI system was performed. MTF of the single-pulse half-resolution mode was measured independently using two different phantoms, the tungsten square traditionally employed in detailed characterisation experiments positioned on the surface of the detector, and the new QEPI1 phantom at isocentre. Taking into account geometrical magnification, both phantoms yielded the same f_{50} of 1.0 cycles / mm (in the plane of the detector) to within 0.1 cycles / mm. This is very encouraging and suggests the QEPI1 may have a role in the optimisation of kilovoltage radiotherapy imaging systems in the future.

Through application of a new phantom it was confirmed that the MTF in the longitudinal direction of a DRR (i.e. the direction of couch motion) deteriorates with increasing slice thickness. This was as expected and is well understood in the radiotherapy community. However, it was also discovered that for very thin slices the MTF begins to deteriorate in the direction parallel to the scan plane. There appears to be an optimum slice thickness of between 1–2 mm which balances the two effects.

Two display devices – a computer monitor and a seminar room projector – were calibrated to conform to the DICOM Greyscale Standard Display Function. When the calibration was checked it was determined that both conformed sufficiently well to be fulfil the role of Class 2 image review workstations. As imaging becomes more and more prevalent in radiotherapy it will be increasingly important for the displays on which the images are interpreted to be regularly tested as part of a routine QA programme.

Throughout this and previous chapters, the flexibility provided by IQWorks has enabled it to be applied both to addressing some real clinical questions, and to exploring assessment approaches for emerging technologies for which there is currently no overall consensus within the radiotherapy community. This flexibility is key to any optimisation strategy and is central to the image evaluation framework described throughout this work.

Chapter 6.

Discussion and Conclusion

6.1. Overview

Throughout the previous chapters a uniform framework for the objective assessment and optimisation of image quality in a radiotherapy environment was introduced. The theoretical framework and the practical approach to implementation and development of IQWorks was documented, and detailed examples were presented of its application to a variety of radiotherapy imaging modalities. Work has also been presented to demonstrate that it can be applied consistently to all imaging modalities across the whole radiotherapy process and, indeed, to wider imaging applications outside of the radiotherapy field.

This chapter evaluates the successfulness of the framework, discusses its reception by other researchers and considers its future scope.

6.2. Evaluation of the IQWorks Framework

A number of factors have led to the explosion in radiotherapy imaging. These include highly integrated radiotherapy information management networks which allow images to be easily transferred between clinical software packages and imaging and treatment devices, the introduction of advanced image guidance technologies — such as high quality EPI, kilovoltage radiographic imaging and cone-beam CT integrated with the treatment gantry — and the ready availability through PACS of radiological studies which can be co-registered with the planning CT scan and then harnessed during treatment planning. Many of these developments have only occurred in the past 5–10 years, with widespread adoption being far more recent.

Ever since the very earliest applications of imaging in radiotherapy there has been recognition that optimisation and quality assurance of imaging devices is

important. However, there has been little consensus about how this should be approached, with techniques varying widely not only between centres but also between the radiotherapy and diagnostic imaging disciplines. Furthermore, until now there has been little attempt to maintain compatibility between the evaluations performed on different modalities, thus preventing the optimisation of modalities against each other throughout the patient pathway. A measure of the successfulness of the framework introduced in this work is whether or not it provides a platform by which this can be achieved. In addition, whether the framework can be implemented into routine practice and used *ad hoc* to address real clinical issues or questions are also important factors in determining whether it achieves its goals.

From a purely scientific standpoint, IQWorks is capable of generating results in agreement with those of other packages, including the gold standards in the field: OBJ-IQ_reduced and DIMOND3 QA-Distri for radiographic imaging and the ImPACT+ evaluation software for CT. As a basic image analysis package it can therefore be considered a success. However, unlike the other packages, the initial design decisions taken to aid integration into a clinical workflow enable IQWorks to be utilised on a routine, daily basis by non-specialised staff. Its 'limits' framework permits advanced analyses to be employed as part of quick quality control checks, generating simple PASS / FAIL results which can be taken at face value without the person performing a check requiring detailed knowledge of the underlying algorithms. Database storage of results enables review — even immediate review from a remote location — by more experienced staff and is invaluable in allowing trends to be identified and tackled pro-actively.

Many commercial phantoms for assessing image quality are underpinned by years of research and development. Analysis software may be available, often as an optional extra, and will be designed to fulfil the needs of users as judged by the phantom developers. Although such software is undoubtedly valuable in bringing a new phantom into use rapidly, providing a solid platform of objective results on which to build, its use longer term or applicability as part of a wider QA programme may be problematic. Analysis algorithms may not yield results which are compatible with similar packages provided by other manufacturers or may follow a methodology not in accordance with current best practice, either in radiotherapy imaging or in the diagnostic sphere. Indeed, because of the closed-source nature of commercial packages it can often be extremely difficult to identify exactly how a particular algorithm has been

implemented, something which is important when measurements produce results not in line with expectation. Examples of some of the limitations of commercial software packages were discussed in chapter 3.

A particular issue with any commercial package is that, having been designed to perform a limited set of tasks under very specific conditions, it will almost certainly not fully meet the requirements of individual end-users. The lack of a feature which is highly desirable to a particular user may be the difference between a phantom being used regularly or not. Forcing images to be analysed in a particular fashion, including potentially the manner in which a phantom is set-up, may not be the most appropriate way to investigate clinical questions. Sometimes, physical customisation of a phantom may be necessary for a particular experiment yet established analysis software will be unable to cope with this. For all these reasons, as the perceived limitations of a commercial phantom and software solution increase, scientists tend to find themselves developing alternative strategies to overcome these and the commercial option falls into disuse. Rather than developing new analysis software from scratch, IQWorks now provides a standard framework on which to build. Being capable of mimicking the performance of established commercial packages it can be used to undertake exactly the same analysis as specified by the phantom designer – if this is so desired – but can also be extended to accommodate new applications, phantom modifications or to generate results compatible with those of other workers. In addition, IQWorks can be employed as an independent analysis tool to validate the performance of commercial packages.

Because IQWorks accepts images in a wide range of file formats, and there are facilities to manipulate the origin of coordinates and pixel size if necessary, it is not limited to a particular modality. In conjunction with its database facilities it therefore provides a central point for the quality control, detailed analysis and reporting of results for almost any modality. Rather than rely on separate packages – all employing different underlying image processing algorithms, reporting mechanisms and storage facilities – to analyse specific phantoms for particular modalities, IQWorks allows all phantoms for all modalities to be analysed using equivalent methods via the same workflow. This goes some way towards breaking down artificial barriers between modalities and disciplines. Furthermore, images from emerging new modalities can immediately be analysed using existing techniques and phantoms, thereby improving optimisation prior to clinical use and encouraging a deeper understanding of their comparative performance.

Optimisation of an imaging modality for a particular clinical application requires an appreciation not just of the requirements of the application but of the technical capabilities of the equipment and the influence of acquisition parameters on the images obtained. Comprehensive optimisation of radiotherapy imaging across the entire patient pathway therefore requires scientists with considerable knowledge and experience. From a technical perspective, IQWorks provides an opportunity for scientists to explore the impact of different acquisition settings using sound objective evaluation techniques. This is particularly important for newer equipment which may not yet be fully documented in the literature. In addition, being able to interactively adjust the behaviour of analysis routines without resorting to programming allows scientists to develop a better understanding of the limitations of both the phantoms they are using and the analysis algorithms themselves. However, because extending the code through programming is also an option there is scope to implement workarounds or enhanced algorithms to address some of these.

One potential criticism of IQWorks is that all possible configuration options for a particular analysis module are presented to the user and it may be difficult to identify what the 'correct' settings should be. However, this is actually one of its strengths because it discourages expert users from considering sophisticated analysis algorithms as simple 'black boxes' which 'just work' but which might in reality be producing fundamentally incorrect results. Through experimentation, reference to the literature and in consultation with other expert users the most appropriate parameters for a given application can be deduced then saved for future use as part of an analysis tree. In this way, parameters are consciously and deliberately chosen to match the application, adding confidence and resilience to the analysis scheme whilst also improving awareness of the underlying science.

Indeed, the flexibility provided by IQWorks will be key to its wider adoption and potential longevity. Imaging scientists who may not be expert programmers can apply standard analysis methodologies to new modalities and phantoms, fine-tuning the algorithms to their requirements. In the examples described in chapters 3 to 5 there were numerous occasions when *ad hoc* adjustments to existing analysis trees were performed based on the results of previous experiments. For example, during the commissioning of the Oxford OBI linacs the analysis tree originally used to analyse images of the Catphan alignment slice was modified to accommodate the RMI electron density inserts taped together in a water phantom. Another example was investigating the potential

of the QEPI1 phantom – originally designed for electronic portal imaging – for use in the QA of kilovoltage radiographic equipment. These experiments added considerable value to the work being undertaken yet were only possible because IQWorks could accommodate changes to the experimental techniques through minimal interactive adjustments.

Another area where IQWorks provides new opportunities is the automation of complex analyses. Traditionally, detailed, time-consuming manual analyses would only be performed during initial equipment commissioning or perhaps for annual QA. However, because IQWorks automates the analysis of even the most complex phantoms, yielding results in only a few seconds, it is possible to perform a whole battery of advanced tests each time a parameter is changed as part of an optimisation experiment. For example, the detailed investigation of EPI MTF across the field of view in chapter 3, and the verification and monitoring of the OBI CBCT electron density calibration curve across all clinical acquisition modes in chapter 4, both involved repetitive measurements which would not have been possible by hand. In each case, IQWorks facilitated a detailed characterisation of the system concerned and indeed, as discovered for the OBI system, such a characterisation may be advisable on a more frequent basis than originally anticipated. In fact, by making complex analyses readily accessible IQWorks makes it viable to incorporate these into frequent PASS / FAIL QC checks.

Being an open-source package, the implementation of algorithms in IQWorks is transparent and open to peer-review, correction, maintenance and quality assurance. Programmers can extend the code to implement their own particular algorithms and any contribution made by an expert individual will benefit the community as a whole. By making the code freely available there is motivation for scientists to adopt IQWorks as the starting point for developing new applications rather than writing new code bases from scratch. Over time, users will potentially benefit from new code additions originally developed to meet very specific individual needs but which may have wider applicability.

One of the original goals of IQWorks was to encourage the exploration and adoption of objective image quality assessment techniques. Through design decisions and the comprehensive implementation of key algorithms IQWorks empowers scientists to focus on the science of their imaging experiment, enabling them to harness state of the art, robust analysis techniques without having to worry about the technical coding and implementation of these. It allows experts to focus on the clinical optimisation task without being overly

concerned with the underlying software implementation, and it is encouraging that users are beginning to harness IQWorks as a platform for exploring how analysis methods or phantoms originally developed for one modality might be applied to others.

As a modality-neutral image analysis package IQWorks has considerable scope for both optimisation and QA applications. It has been incorporated into the routine workflow of a number of radiotherapy and diagnostic imaging departments and underpins the radiotherapy imaging QA programme at the new Oxford Cancer Centre.

6.3. IQWorks and the Community

In 2007, IQWorks was selected by an IPEM consensus group to form the basis of a national software package to champion objective image quality evaluation. Since then there has been widespread support for the project from both the radiotherapy and diagnostic imaging communities. A website has been established (http://www.iqworks.org) to manage documentation, user support and official code releases and this is rapidly growing into a fertile collaboration platform. Users are increasingly sharing sample phantom images, analysis trees they have constructed and improvements or fixes to the code they have developed.

At present there are over 400 registered users, with there being representatives from every continent of the world. In addition, there has been considerable support and encouragement from technology user groups and professional bodies including the CT User Group (CTUG), Digital Radiology User Group (DRUG), Institute of Physics and Engineering in Medicine (IPEM) and the British Institute of Radiology (BIR).

In October 2009 IQWorks was the subject of an oversubscribed national scientific meeting and a second meeting concentrating on code development is planned for early 2011. IQWorks has also been the focus of over twenty conference papers, some of which are listed in Appendix D.

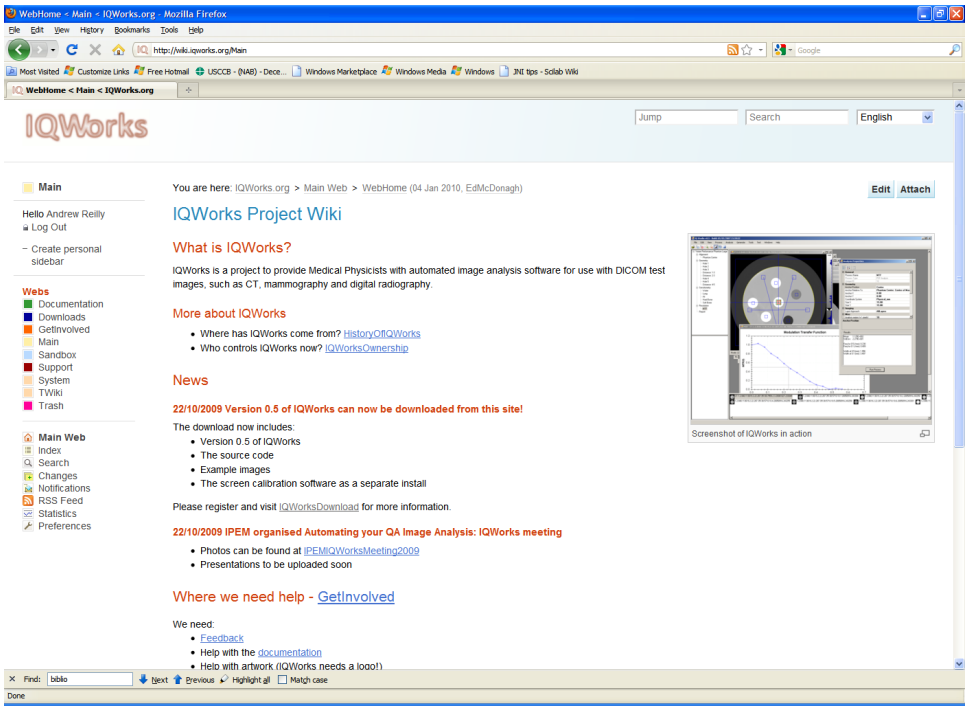


Figure 6.1. – Screenshot from IQWorks website.

6.4. Future Work

6.4.1. Expansion of Scope and Utilisation of IQWorks

An immediate priority is to raise awareness of the potential of IQWorks, encouraging its wider adoption in both radiotherapy and diagnostic imaging via the above 2010 and subsequent meetings, further presentations and via publications based on this work (see list of possible titles in Appendix D). One of the experiences of the author has been that there is sometimes a perception that IQWorks is intended for a specific application and therefore not suitable for others. Work is therefore required to demonstrate its applicability across modalities and disciplines as a useful tool for routine quality assurance, clinical optimisation exercises and also for more fundamental research and development. A long term goal is to support collaboration and convergence between clinical and scientific specialities by establishing the acceptance of robust and comparable objective image quality metrics between different modalities.

With the initial phases of development complete, greater impetus will be put behind the open-source IQWorks project. There has been considerable interest from the community in contributing to all aspects of this, including code validation and quality assurance, exploring the potential use of the system

with new applications, phantoms and modalities, and writing documentation. To be successful in the long term, such a large collaborative effort will require a significant amount of organisation and management. Regular user group meetings will also be held to provide an opportunity for novice users to meet with experts and forge collaborative links.

A particular aspiration is to encourage respected experts to take responsibility for maintaining and developing elements of the code base within their realm of expertise. There is motivation to invest effort in this because it encourages widespread adoption of new methods as a whole and also makes IQWorks viable as a package to be referred to in national standards and professional guidance. Ultimately, the goal is for IQWorks to be a reference implementation of best-practice analysis methodologies.

6.4.2. Development of Phantoms and Support for Phantom Use in Conjunction with IQWorks

Two novel phantoms were developed as part of this work – the QEPI1 phantom and the DRR geometry phantom – and a number of existing phantoms were modified or improvised to accommodate particular applications. It is intended to commercialise some of these efforts and it is envisaged that IQWorks and phantom development will progress in tandem in the future, each supporting the other. Currently, there is no commercial phantom for CBCT which occupies a sufficiently large field of view, contains electron density inserts made from biologically representative tissue-equivalent materials and which also incorporates test objects for overall image quality assessment (such as an MTF bead or angled ramps to consider slice thickness). There is therefore a significant opportunity to design a comprehensive new 'all-in-one' phantom for CBCT QA and calibration.

In terms of support for currently available phantoms it is intended that a library of analysis trees for different phantom / modality combinations be developed and made available for download from the website. This will be achieved through a combination of pro-active development and encouraging contributions from the community. Certainly, analysis trees will be developed to meet the requirements of current professional guidance, but it is also hoped that the writers of new guidance will develop reference analysis trees to complement their publications. Ideally, an analysis tree would be available for every commercial phantom, with the vendors aware of this so that they can sell their

equipment as 'compatible' or 'supported' by IQWorks. In some cases, where there is no commercial conflict, vendors may even promote the use of IQWorks with their equipment. There is also scope to collaborate with the modality manufacturers themselves to develop better phantoms and analysis routines to assist users in harnessing the full potential of their systems, possibly with a view to utilising IQWorks during acceptance testing and routine QA.

6.4.3. Radiotherapy Imaging Intercomparison and Audit based on IQWorks Consistent Approach

Throughout 2010 the IPEM South East Dosimetry Audit Group will be embarking upon a programme of intercomparisons of the performance of radiotherapy imaging equipment. This will cover image quality and dosimetry aspects of CT, CBCT and EPI and it is envisaged this pilot study will evolve into a national UK audit in 2011. All image quality evaluation will be performed using IQWorks and the QEPI1 phantom will underpin the EPI measurements. By providing comparative performance information these audits will help to standardise imaging techniques across the UK radiotherapy community, as well as promote the application of objective assessment methodologies in radiotherapy imaging.

6.4.4. IQWorks to Provide a Uniform Approach to Display Device Assessment and to Underpin Human Observer Models

QA and calibration of display devices is a relatively new area of interest in radiotherapy and IQWorks makes viable assessments which previously were prohibitively expensive or time-consuming. It therefore has real potential for raising the profile of this topic and being key to the development of a pragmatic approach that meets the requirements of radiotherapy clinical applications. Furthermore, with the widespread availability of PACS images through standard web browsers running on consumer-grade computing equipment there is also considerable scope for this application of IQWorks in the diagnostic sphere.

Finally, the ultimate goal is to develop a mathematical model of the human observer performing clinical radiotherapy tasks, as other researchers have had some success in achieving for diagnostic imaging. This is significantly aided by IQWorks providing a uniform framework for evaluation of all radiotherapy imaging modalities, and also by it providing a consistent evaluation of the

ubiquitous final link in the human-to-imaging system interface, i.e. a human interpreting images on a display device. This will facilitate a deeper understanding of the role of each imaging modality as part of cancer management, thus aiding in the optimisation of each individual modality and of the radiotherapy process as a whole.

6.5. Conclusion

This work presented a flexible conceptual framework for objective image quality evaluation which has been implemented practically through the IQWorks software package. Its immediate applicability to a wide range of radiotherapy imaging modalities has been demonstrated and opportunities for optimisation have been discussed in the context of real clinical questions which the new framework has helped address. It is clear that the software and supporting phantoms have considerable scope for development in the future, not just in radiotherapy but in medical imaging as a whole.

IQWorks has been positively received by both the radiotherapy and diagnostic imaging communities and it is hoped this work will continue to contribute to the adoption of objective measurement techniques by both disciplines in the future, as well as improved collaboration and convergence between the two.

Appendix A.

IQWorks 1 — A Framework for Image Quality Evaluation

A.1. Overview

Chapter 2 introduced quantitative techniques for objective image quality evaluation. However, the power and intrinsic benefits of advanced analysis algorithms can only be realised in a regular hospital environment if standardised tools are available for loading, interpreting and processing images. Furthermore, if these are to be incorporated into a routine QA programme the tools must be quick and straightforward to use, and provide appropriate functionality for the reporting and storage of results. To date, there has been no readily available package providing these features comprehensively across all modalities. As a result, detailed quantitative image quality measurements have tended to be restricted either to a laboratory scenario, or to centres where individuals have developed in-house code implementing specific algorithms for particular applications. In addition, with advanced imaging modalities only recently becoming widely available to radiotherapy departments, there is generally insufficient time for physicists to develop fundamentally new measurement techniques alongside the clinical implementation of the new equipment. Despite recommendations regarding quantitative image quality evaluation having existed for some time for most diagnostic imaging modalities – and to some extent those in radiotherapy – it would be true to say that at present these are only sparsely implemented.

IQWorks is a general purpose package for the objective, quantitative and automated evaluation of image quality and it forms the foundation of this work. It has the primary goal of enabling the routine application of sophisticated image quality evaluation methods in a standard hospital environment, across

all modalities in both radiotherapy and diagnostic imaging. IQWorks has been developed by the author in Edinburgh and Oxford over the course of eight years and consists of some 30,000 lines of computer code.

This appendix presents the rationale behind the IQWorks system, the underpinning philosophy of a comprehensive QA tool and introduces the concept of the 'universal phantom'. This is a new approach which enables IQWorks to potentially be directly applied to the analysis of any phantom imaged on any modality, and is a route to improving consistency of approach across disciplines and modalities. Key development goals are considered and the implementation of these discussed in terms of design decisions, software development, image handling, algorithmic development, and the reporting and storage of results. Validation and QA of the package are also considered. Other software frameworks supporting quantitative image analysis also exist and these are compared against IQWorks, illustrating its strengths and limitations, and also opportunities for synergy.

Flexible numerical implementations of algorithms for calculating the metrics discussed in chapter 2 have been developed to be available within the overarching IQWorks analysis framework. These are described in detail in Appendix B.

A.2. Introduction to IQWorks

Calculation of advanced image quality metrics requires numerical implementations of complex evaluation algorithms and the processing of phantom images by computer. For as long as it has been possible to digitise analogue images, and especially since the advent of the television tube, scientists have been developing computer programs for image analysis and working to apply these to the whole spectrum of imaging modalities. It has long been a goal of most medical physics departments to incorporate such routines into their QA and commissioning programmes, and the majority of departments will have developed some in-house code to achieve this for specific projects. Commercial packages have also gradually become available to analyse particular phantoms imaged on a specific modality, and more recently general purpose tools which are applicable across a range of phantoms and modalities.

In 2001, the 'Phantom Analyser' package was created by the author as an in-house analysis tool for CT images. This, the predecessor to IQWorks, was a far more simplistic package designed to perform a limited range of analyses

on a small number of CT performance phantoms. It was intended to support the Edinburgh Cancer Centre QA programme for the new CT Simulator and Sim-CT systems installed at that time, and was successfully used for this over a number of years. As clinical applications became more sophisticated, and the requirements of imaging more demanding, it quickly became desirable to extend the range of phantoms supported by the package, and also to be able to analyse images from other modalities, particularly EPI, DRR and MRI. However, because the analysis routines for different phantoms were hard coded, and in addition incorporated 'fine-tuning' or caveats appropriate to one modality but not others, the more general the requirements of the software, the more difficult it became to extend it. Before long, the effort required for ongoing development became greater than the benefits realised from the time investment required.

This is a common theme with in-house software packages, and one which inhibits the wider applicability and limits the longevity of many tools developed in hospital departments today. Because most in-house programs tend to be written, often quickly and expertly, to meet the needs of a specific application, it can be extremely difficult to modify them to meet the needs of others. Furthermore, the most flexible in-house tools are usually also those where considerable effort has gone into the coded implementation of the core analysis algorithms, but little towards its user-friendliness or extensibility. Unfortunately, this results in tools which are straightforward to use in the hands of a developer, and in his or her department, but which are not easily transferable to be used by others. In a similar vein, it is frequently the case that extremely useful code written by one individual, and which may have been relied upon for a number of years, suddenly becomes impossible to maintain when the individual leaves. For these reasons, in-house image analysis packages are frequently written and rewritten from scratch, often resulting in largely the same functionality as was available previously.

When Phantom Analyser was first being developed, a number of commercial packages were also available for image quality assessment, and two of these were evaluated in Edinburgh: PIPSPro (Masthead Imaging Corp, British Columbia, Canada) for EPID images of the QC3V portal imaging phantom, and AutoQALite (IRIS Inc, Maryland, USA) for CT images of the Catphan series of phantoms. Whilst both were quick and reasonably robust when used exactly as the manufacturer intended, a number of limitations quickly became apparent. Firstly, only the specific phantoms intended for use with the packages were supported, and even those only when the phantoms and imaging protocols

were set-up as expected by the software (field of view, reconstruction filters, etc.). This meant that the use of the packages for detailed image evaluation experiments, possibly involving in-house modifications to the standard phantoms, or degrading images to mimic particular clinical scenarios, was impossible. In addition, and of more concern, was the fundamentally different and incompatible way in which supposedly 'equivalent' image quality metrics, particularly MTF, were calculated by the two packages, thus preventing performance comparisons from being drawn between the different modalities. Lastly, although the automated nature of the AutoQALite package made it extremely efficient to use, this also resulted in it being difficult to identify the steps involved in the analysis algorithms, effectively reducing the package to a 'black-box' employing trusted yet not completely specified analysis algorithms. For all these reasons, there was a strong motivation to abandon the use of the commercial packages for all but the most routine measurements and to rely instead on in-house code for developmental and more advanced work. This is a recurring theme in many hospital departments, where commercial solutions for scientific applications only partially meet user requirements, leaving a void which must be filled through improvised in-house development.

Nevertheless, a significant benefit of the commercial systems was that both included functionality for the recording and review of results, recognising the importance of these for a routine QA programme. However, the way in which this was implemented was different between the two, and relatively regimented so that the QA programme had to fit around the approach of the software, not the other way round. Again, there was significant motivation to develop software at least to extend the commercial offerings to present results in a fashion more in-line with the needs of the department.

In early 2003 it was decided to rewrite the Phantom Analyser package following an entirely new philosophy. The new package, IQWorks, was intended to be a robust, user-friendly and extendable image analysis platform suitable for both routine and developmental applications, and which overcame all of the limitations outlined above.

Central to IQWorks is the concept of the *universal phantom*: regardless of the specific imaging modality being considered, or the details of the image acquisition protocol, all physical phantoms comprise test objects, arranged in a particular geometrical configuration, for performing a subset of a finite range of elemental image performance evaluation tasks. For example, a phantom might contain an impulse object, line or edge for MTF evaluation; uniform patches for

investigating CNR or uniformity; or high contrast structures between which distances are measured to verify geometric linearity. Any individual phantom can be modelled in IQWorks by constructing an *analysis tree* consisting of *analysis modules* incorporating the image processing algorithms required for the appropriate elemental evaluation tasks, fine-tuned for the specifics of the modality, phantom and acquisition protocol. Phantom images are then analysed in a fully automated fashion by running the appropriate analysis tree, with results being displayed to the end user or recorded as necessary.

Each analysis module is designed to handle a particular elemental evaluation task, as detailed in the relevant literature. Inputs to a module allow the user to make choices regarding the implementation of an algorithm, or to provide the algorithm with fundamental operating parameters. All modules are also capable of outputting the results of their calculations, as well as error and status information. The intention is for any given module to provide a high-quality, robust and self-contained implementation of a single analysis algorithm, aiming to retain maximum utility and flexibility through judicious provision of inputs and outputs and without requiring additional programming. Over twenty elemental analysis modules have been implemented so far in IQWorks and the principal modules are described in detail in appendix B.

A key feature of IQWorks is that a uniform framework is employed to describe the inputs and outputs of all analysis modules. This means that the output of any module in an analysis tree can be used as input to another module further down the tree. Furthermore, standardisation of module inputs allows straightforward and interactive customisation of parameters by the end user via a generic graphical user interface. This is available immediately even for completely new modules, without resorting to additional interface programming. Complex IQWorks analysis trees can therefore be constructed by the user interactively, with the parameters of each analysis module being iteratively adjusted and tested as required. Within an analysis tree the modules can be organised in nested *analysis groups*, enabling the different stages of the image processing scheme to be logically separated. Although these groups influence the processing order of modules and the presentation of results, they do not affect the calculation algorithms or results themselves.

Once a tree is complete, it can be saved to disk for loading again and running when required. Existing analysis trees can also be modified and extended as necessary, giving the user considerable flexibility. For example, basic analysis trees prepared for a routine QA programme can very quickly be extended to

generate additional output for in-depth development applications. Alternatively, if a commercial phantom has been customised as part of an experiment, or a new phantom is being imaged on a different modality, it is straightforward to select an existing tree then fine-tune it to match the new scenario. IQWorks therefore enables sophisticated analysis schemes to be rapidly developed using standardised algorithms and extended as required to accommodate new phantoms, modalities and applications.

A simple IQWorks analysis tree is illustrated schematically in figure A.1. This tree has been constructed to analyse CT images of the alignment module of the Catphan 504 phantom (Phantom Lab, New York, USA), commonly bundled with radiotherapy cone-beam CT systems. On the left hand side in the figure the alignment tree is presented as displayed in IQWorks. It is divided into four groups, each containing a number of processing modules. When the tree is run the modules are executed sequentially from top to bottom in the tree. The underlying logic and linkages within the tree are described here, with technical and implementation details of the modules discussed in appendix B.

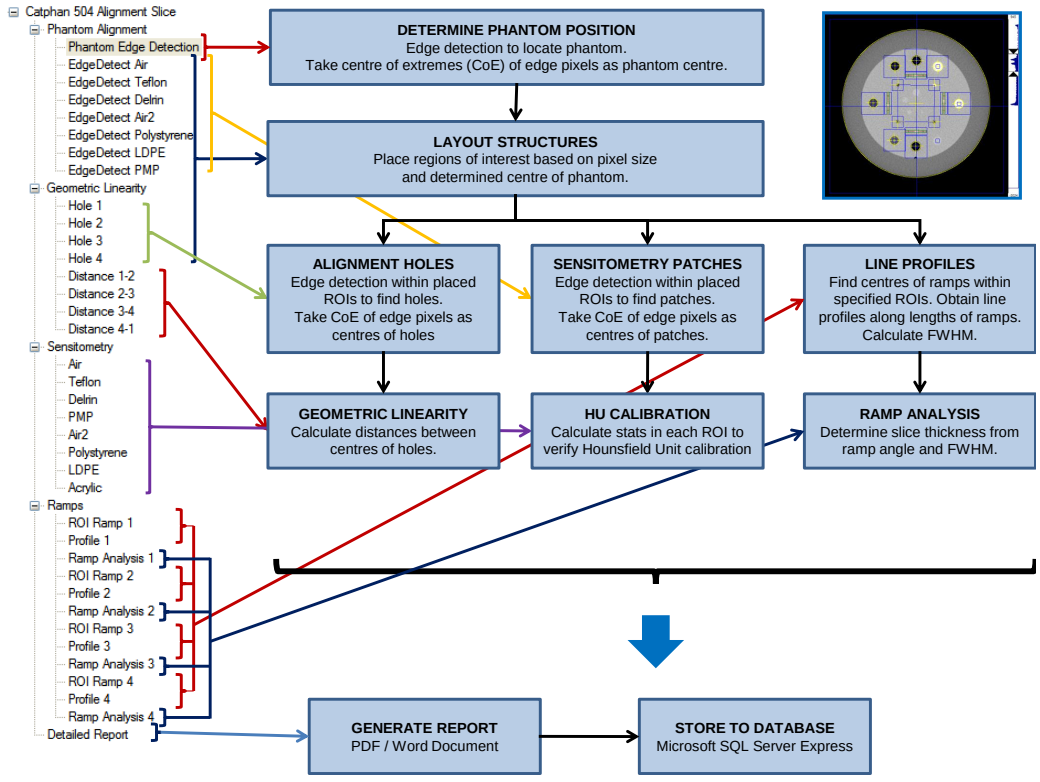


Figure A.1. – Schematic diagram illustrating an IQWorks analysis tree. This tree is used to analyse CT and CBCT images of the alignment module of the Catphan 504 phantom (CT image shown top right).

It is assumed that the phantom may be positioned anywhere in the field of

view, and the first group, 'Phantom Alignment', is concerned with localising the phantom and place ROIs for analyses further down the tree. All the modules with the 'Phantom Alignment' group are actually instances of the same basic module, fine-tuned for the particular object they are intended to localise and labelled accordingly. The 'Phantom Edge Detection' module places a rectangular ROI centred on the origin and encompassing 95% of the FOV. An edge detection algorithm is run within this region and the largest continuous outline within the ROI returned as a series of coordinates. It is assumed this contour represents the outer edge of the Catphan phantom. From the contour data, the extreme values of the edge in the X and Y directions are determined, the width and height of the phantom calculated and the mean of the extremes being taken as the centre of the phantom. Pixel size can be calculated from the dimension information and the centre of the phantom is representative of the accuracy of alignment lasers. All other ROIs are placed relative to the calculated centre of the phantom. (i.e. If the phantom was physically offset 1 cm to the left hand side of the origin of the imaging device, the calculated centre of the phantom should also be offset by 1 cm in this direction and all ROIs for processing modules later in the tree would automatically be shifted accordingly.)

Because the sensitometry patches for HU calibration are relatively small, it is important the ROIs used in the sensitometry analysis are accurately aligned and a gross alignment correction for the whole phantom may not be sufficiently accurate for this. Therefore, further edge detection is performed within smaller regions placed around the expected positions of the sensitometry patches. In a similar manner to the gross phantom, these are used to localise the patches more precisely and place ROIs accurately at their centres.

Intersecting the alignment module are four cylindrical pins, located at the corners of a square of side 5 cm. These are provided to assess geometric linearity and appear as circular holes in the image, 3 dark and 1 bright. In the 'Geometric Linearity' group further edge detections are first performed to localise the holes and determine their centres. Distances between the centres are then calculated, from which pixel size can be determined and geometric linearity in the two pixel matrix directions assessed.

Simple statistical calculations are performed by the ROIs placed in the 'Sensitometry' group, with the mean Hounsfield Units in each region either being compared against the expected values for the given scanner, or being used to determine the characteristic curve for a new system by plotting HU against the known electron densities of the different sensitometry materials. The standard

deviation in HU in each ROI is related to the stochastic noise in the image and should therefore be relatively constant for any given scanner at a particular operating point.

Interpretation of the four angled ramps intersecting the alignment module is a three step process and performed by the 'Ramps' group. For each ramp a region of interest is placed around the expected location of the ramp, the centre of the ramp is identified as the maximum pixel value within the ROI, and a line profile calculated about this. The average of 3 adjacent profiles is taken each time to minimise the influence of noise (centre profile ±1 profile either side). From the full-width of the line profile at half its maximum intensity (the FWHM), and the known 23° angle of the ramps, the effective imaged slice thickness can be calculated. In addition, the relative locations of the centre of the FWHM and the centre of the phantom can be used to determine the location of the slice in the direction normal to the plane of the slice.

At the end of the tree is a 'Detailed Report' module which generates a formatted PDF or Microsoft Word document summarising the results of the analyses performed, presents this to the user for approval, then stores the details of the results into an SQL server database for future retrieval and further analysis. This module also allows the results of measurements to be compared against expected baselines, presenting a 'pass' or 'fail' result to the user. The comprehensive reporting and database storage functionality of IQWorks is described in more detail in sections A.9 and A.10 below.

Although the analysis tree described above was developed initially to be applied to kilovoltage CBCT images, it can immediately be run against images from other x-ray CT modalities – including conventional CT scanners and megavoltage CT and CBCT systems – without requiring any modification. It can also be applied to CT images of other phantoms in the Catphan series of phantoms with only minimal changes. For example, the Catphan 600 alignment module is identical to that in the Catphan 504, except that it is rotated through 180° and the sensitometry object labelled 'Air2' in the tree is filled with water. Adjusting the analysis tree is extremely straightforward, requiring only the relabelling of entries and updating the expected pixel values within the new water region. Similarly, the tree can easily be adjusted to accommodate the alignment module of the Catphan 500 phantom (by deleting entries from the Phantom Alignment and Sensitometry groups to take into account the fewer sensitometry objects) or the equivalent module in the AAPM Performance Phantom (CIRS Inc, Virginia, USA).

Analysis trees for a number of common CT phantoms are described in chapter 4, those for EPI in chapter 3 and some examples for other modalities in chapter 5.

IQWorks is aimed at meeting the demands of users working with any phantom imaged on any modality. In general, there are three categories of user: those writing fundamentally new analysis algorithms to accommodate an elemental analysis module not currently provided by IQWorks; those constructing analysis trees to perform experiments or underpin a routine QA programme; those running analysis trees against phantom images acquired as part of a routine QA programme. When developing a new analysis algorithm it is necessary to write computer code to augment the IQWorks framework. As described in section A.4 below, this code can either be part of the IQWorks package itself or form the basis of an external module which is launched by framework. However, as long as suitable analysis modules are available, all analysis tree development, data reporting and running of trees against images can be performed via the graphical user interface, even for very advanced or complex applications.

A significant amount of effort was expended developing the IQWorks user interface, with the rationale being that a clumsy or obstructive interface could be the difference between the package being widely used or abandoned. In particular, an efficient, user-friendly interface is absolutely crucial in any tool utilised for routine QA. Furthermore, enabling all analysis modules to be manipulated and fine-tuned via a common interface helps the scientist to concentrate on the science and algorithms underpinning an analysis tree, rather than being overly concerned with the technical implementation of the analysis components themselves. This facilitates the development of sophisticated analysis schemes for modalities such as EPI and DRRs where previously there has been only limited work in this field.

Figure A.2 contains a screenshot of a typical IQWorks session and illustrates the analysis of a CT scan of the Catphan 504 alignment slice using the analysis tree described above. Key features of the user interface are labelled in the figure.

At the top of the IQWorks screen is a series of menus which gives access to all functionality provided by the package, including loading and saving images; visualisation tools; constructing, saving, loading and running analysis trees; and ad hoc image processing such as inverting or rotating images. Immediately below the menu bar is a strip of *accelerator buttons*, graphical icons which provide shortcuts to the most common options contained within the more complex

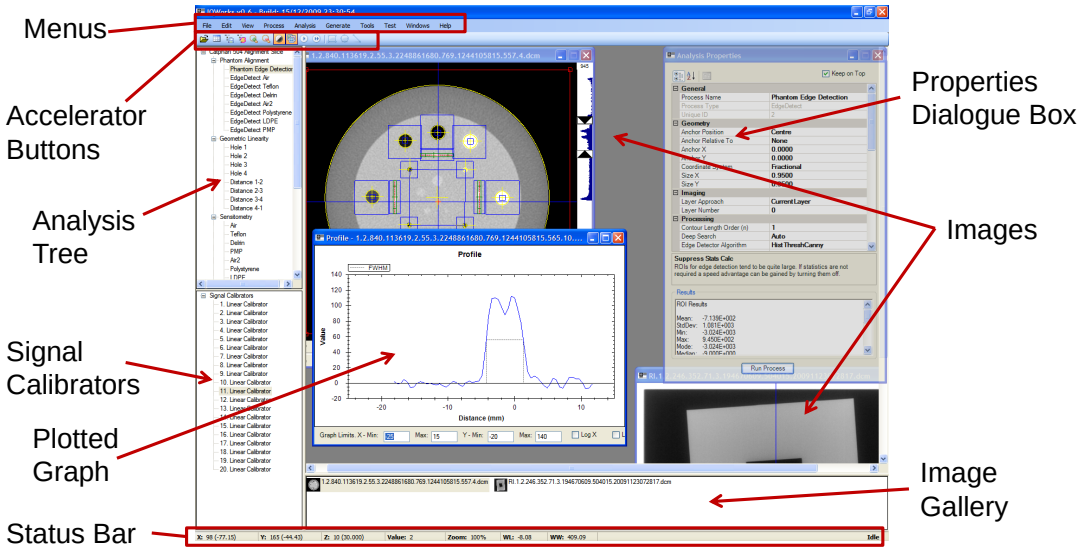


Figure A.2. – Screenshot illustrating the key features of the IQWorks graphical user interface.

menu structure. These provide quick access to methods for loading images; loading and running analysis trees; visualisation tools such as zooming and windowing; the placement of rectangular or elliptical regions of interest; and distance measurements. If the mouse hovers over a particular accelerator button a small 'tool-tip' is displayed containing a brief description of the functionality provided by that button.

When an image is loaded it is displayed in the main image pane, and added to the *image gallery* at the bottom of the screen, which provides a thumbnail of the image alongside its filename. The origin of coordinates of an image is indicated by dark blue cross-hairs extending across the field of view. As many images can be loaded as will fit into the memory of the computer, which can number into the hundreds for modalities with relatively small image matrix sizes, such as CT or EPI. The gallery provides a convenient means of locating and selecting particular images amongst those loaded and for selecting multiple images simultaneously. When an image is selected its thumbnail and filename are highlighted in the gallery. In figure A.2, two image sets are loaded: a CT scan of the Catphan 504 phantom, with an image of the alignment slice currently being displayed, and an EPI image of the QEPI1 phantom. The Catphan 504 image set is currently selected.

Analysis trees are constructed against and thus 'belong' to a particular image instance, although they can be loaded or copied and pasted onto another. When any given image is selected, the analysis tree attached to that image is displayed

in the top-left panel in the user interface. Analysis modules can then be added to or removed from the tree via the menus and accelerator buttons. As a tree is being constructed, individual analysis modules or whole processing groups (effectively the branches of the tree) can be reordered by dragging and dropping with the mouse, or duplicated by copying and pasting elsewhere in the tree. Many analysis modules have a spatial manifestation, such as a region of interest in the image or a line profile between specific points. Graphical representations of these are drawn on top of the image itself, along with additional annotations when an analysis module has been executed. Furthermore, when an analysis module is selected so that it is active, its structures are highlighted in the image. For example, the 'Phantom Edge Detection' module is currently selected in figure A.2 so that its search ROI is indicated by the highlighted red rectangle on the Catphan 504 image. In the figure this module has already been executed, with the detected edge being shown by the yellow circle and the centres of the extremes of the edge in the X and Y directions (i.e. the detected centre of the phantom) by the large yellow cross.

All operating parameters for individual analysis modules can be adjusted interactively via the *Properties Dialogue Box*. This provides access to module inputs and outputs via a common interface, and also a summary of results when a module has been executed. Numerical or textual parameters can be modified by typing in the new values, after which entry validation is automatically performed to ensure the values are of the correct type and range. Other options can be set by making selections from drop down lists. Where an output from another module is expected (such as the location of the centre of a phantom, against which ROIs will be placed) then all appropriate outputs from modules earlier in the tree are also presented to the user in a drop down list. Both the assignment of operating parameters, and the forming of links between modules, are therefore straightforward processes where the user interface ensures only valid choices are made. This avoids problems when trees are later run in their entirety, thus facilitating automated processing of multiple images. To aid fine-tuning of parameters each module can be executed individually via the properties dialogue box, and where possible the effects of any changes made to parameters are displayed immediately. As well as via the properties dialogue box, some modules also allow interactive manipulation of their parameters by the mouse on the surface of the image, such as the placement and resizing of regions of interest.

In chapter 2, the importance of image data being linear for many analysis

routines was discussed. If the pixel values in an image are not inherently linear then they must be made so before performing image quality calculations. IQWorks *Signal Calibrators* perform this linearisation step, providing a mapping between a raw pixel value and its calibrated, linearised value. It is the image of calibrated pixel values which is passed to IQWorks analysis modules and displayed in the user interface. Each image plane is assigned an independent signal calibrator, which by default performs a direct mapping with no modification of pixel values. However, a number of possible calibrators exists, each with their own parameters that can be manipulated via the user interface, and the different planes of an image need not have the same type of calibrator, or calibrators with the same parameters. In figure A.2, the signal calibrators corresponding to the twenty CT slices in the scan of the Catphan 504 are displayed in the bottom-left panel of the interface. Signal calibrators are discussed in detail in section A.8 below.

At the very bottom of the IQWorks screen is a *status bar* providing real-time information about the state of the package. As the mouse moves across an image, the coordinates of the point under the cursor location are displayed simultaneously in image and real-world (physical) coordinates, both in the plane of the image and orthogonal to this plane (i.e. up and down a stack of images). Coordinate systems are considered further in section A.7. The calibrated value of the pixel under the cursor is also automatically updated as the mouse moves. Additional status information includes the zoom level and windowing settings of the active image, and whether the package is currently calculating or awaiting user interaction.

When analysis modules are executed, either individually or as part of a tree, graphical results will sometimes be generated, such as additional images, plotted graphs or surface renderings. These are displayed alongside other images in the IQWorks user interface, with new images that can themselves be analysed using analysis trees being also added to the gallery.

Attainment of IQWorks' primary goal of providing a uniform analysis framework applicable to all phantoms across all imaging modalities will only be possible if there is sufficient scope for user customisation. As imaging science advances, new and refined analysis algorithms will regularly become available which have not yet been implemented in IQWorks. Rather than being a barrier to its uptake, it is hoped that IQWorks will provide a mature platform for developing and exploring these algorithms, with it then being straightforward to deploy them for routine applications once ready. The aspiration is that experts

in different fields will contribute code, algorithms and expertise to encourage the adoption of their methods and facilitate standardisation across imaging modalities and disciplines. Ultimately, it is hoped that IQWorks will be adopted as a standard analysis tool that, with reference analysis trees, is cited in the professional guidance literature for specific modalities and applications.

Within the medical imaging community there is a strong will to support activities which benefit the whole community, but a wariness of the inflexibility and potential conflict of interest of commercial offerings. It was therefore decided early on to release IQWorks as an open-source project, with all the source-code being available for free under the GNU General Public Licence, more details of which are provided in appendix F.1. Essentially this licensing ensures individual contributors retain the copyright to their work, and that the code can continue to be used for any purpose as long as the source-code for any derivative works continues to be freely available. Aside from making the package readily available to as wide a potential user-base as possible, releasing the IQWorks code open-source has a number of advantages. As uptake increases, there will be strong motivation and support to develop new modules to work within the existing framework, and to refine the framework itself, rather than continually starting from scratch, which has tended to be the case previously. Furthermore, scientists in the field will be able to identify exactly how algorithms are numerically implemented, thereby understanding their limitations, and students can use the code as reference implementations of standard algorithms. With no commercial bias, there is also the potential for IQWorks to be referenced by standards and guidance documents. Finally, after only a short period of time the quality assurance intrinsic to an open-source, multi-contributor project will result in the entire IQWorks code-base essentially being 'peer-reviewed' so that its strengths, limitations and general performance will be generally acknowledged and accepted. Code QA is discussed further in section A.5.

IQWorks was positively received when it was first made available to the community. At a meeting held jointly by the UK CT User Group and IPEM's Diagnostic Radiology Special Interest Group in 2007 it was unanimously decided to adopt IQWorks as the preferred automated image quality evaluation package for the UK. Since then IQWorks has been widely adopted both in diagnostic imaging and radiotherapy, and there is strong interest in contributing to the project. The first scientific meeting concentrating on IQWorks as an analysis package was held in November 2009, with a follow-up developers meeting

planned for late summer 2010. There is clear support from the community, providing confidence that the IQWorks framework has the potential to meet its primary and wider goals.

Over the past six years IQWorks has formed the basis of an increasing number of elements of the imaging QA and development programmes in the radiotherapy departments in both Edinburgh and Oxford. Some examples of IQWorks applied in these are provided in chapters 3–5.

A.3. Design Decisions

In aiming for IQWorks to be as widely applicable as possible, and to encourage uptake, it was very important to ensure the package was robust, reliable and user-friendly. As described above, it was essential that consideration was given to both the user interface and numerical implementations of the underlying algorithms, with each aspect of the package being important to different users. To facilitate development of algorithms without requiring frequent reprogramming of the interface it was decided to separate the two as much as possible. Standard programming structures were developed through which all analysis modules communicated with the user interface. This allowed either the user interface or any individual analysis module to be revised without any risk of the changes having major repercussions elsewhere in the code base. Furthermore, each analysis module has been written to operate independently of the others, with inter-module communication again being through standardised programming structures. In this way, once a module has been developed and tested it can be assumed that it will be compatible with modified and new modules in the future, with its algorithmic implementation only requiring thorough testing again if changes are made to the underlying code of the module itself.

Object-orientated design (OOD) was employed to structure the IQWorks code into modules which were as isolated and independent as possible. Under this methodology all functional units are considered an *object*, including the analysis modules, the parameters input to a module, the image the module works with, etc. A *class* describes a particular type of object and has associated data and *methods* which can be used to act on that data. For example, a region of interest may be a class which has a 'calculate statistics' method. A new class can *inherit* functionality from a base class, extending this, but also retaining compatibility with other classes and methods which were originally designed to interact with the base class. Using the ROI example again, IQWorks contains

analysis modules for rectangular and elliptical ROIs, as well as annular versions of these. These all extend from the base 'PROI' class which provides basic methods for calculating statistics, user interaction, etc. Although the sub-classes all provide different functionality, and are separate classes, they can all be used interchangeably whenever a PROI class is expected. A major advantage of OOD is that individual modules of code can be designed to be completely independent, exposing a documented set of interfaces for interacting with them. If the code is subsequently modified, as long as the interfaces are the same, then they should be able to plug-in to existing code without further modification. Furthermore, automatic code re-use through inheritance ensures sections of code only need to be thoroughly tested once. OOD is therefore a robust, reliable approach to developing mission critical software. Although the software design and management overhead with OOD may initially be higher than in traditional scientific programming, the modular approach results in debugging and long-term development requiring considerably less effort.

Another consideration when developing IQWorks was the hardware platform and operating system under which it would run. In a modern healthcare environment, and particularly a networked radiotherapy department, many computers will already be found in each functional area of the department. Rather than introduce another computer specifically for running IQWorks, it would be attractive if IQWorks could run alongside software on existing computers. Where computers are provided for basic secretarial functions, such as word processing, checking an electronic diary or updating spreadsheets, this is viable with little effort, except that administration access may not be available except on a one-off case to perform an initial installation. However, where the computers available are part of a designated clinical system, and thus classified as 'medical devices' considerable care must be taken. Therefore, ideally IQWorks would not require a formal installation and would also run in isolation to other packages on a computer so that there was no possibility of interference with a medical device.

Currently, the vast majority of computers in a hospital environment are IBM compatible machines (either 32 or 64 bit) running a version of Microsoft Windows (usually Windows XP, but sometimes Vista and occasionally Windows 7). It was therefore decided to ensure that IQWorks, at the very least, would run successfully with full functionality on a Windows platform. However, in the medical imaging community there is considerable support for Linux and Unix based operating systems, particularly for server platforms, so there was also a

will for IQWorks to run under these. One potential future development route for IQWorks is for the user to construct analysis trees using the interactive Windows user interface, then to start the system running in the background on a Linux server that automatically accepts images, processes them with the appropriate tree, and outputs the results to a directory on disk or a database with no user interaction. Another possibility is for analysis trees and images to be uploaded via a web-interface, with IQWorks possibly running on Linux / Apache Server platform. Either of these evolutions would be straightforward technically to implement, but both require portions of IQWorks to be cross-platform.

Another common scenario, especially in diagnostic imaging, is for a physics team to attend at a modality with a laptop dedicated to image analysis or data recording. Alternatively, images may be stored in a PACS system then retrieved for analysis at a standard desktop computer. IQWorks must therefore be capable of performing the whole range of image evaluation operations on relatively low specification hardware, whilst also maintaining sufficient flexibility to be capable of working through many images automatically as part of more advanced research or development work.

Different imaging modalities store images in inherently different formats. Although most modalities support the export of data according to the DICOM standard, implementations vary between different manufacturers and even between systems from the same manufacturer. Furthermore, some disciplines traditionally rely on particular image file formats, even though these have been superseded. For example, in nuclear medicine there is still a tendency to work with Interfile, rather than DICOM images. If additional steps are required to convert images into a format suitable for use with IQWorks then the attractiveness of the package is reduced, particularly as an efficient tool for routine QA. It was therefore decided that IQWorks should support as wide a range of implementations of the DICOM standard as possible, as well as other commonly encountered file formats, including raw binary and text files. Depending upon the nature of the modality being considered, image data can be transferred via a number of routes. For example, images are commonly copied onto an external medium, such as a memory stick or CDROM; alternatively, they might be sent across the network through a DICOM link with PACS; or they may be placed on a shared network drive. IQWorks must be capable of accessing images via any of these routes, including being able to intelligently handle shared data stores containing many thousands of images, possibly with cryptic filenames. Recognising that there will always be certain types of analysis

not possible in IQWorks, provision has also been included for writing out image data in a range of common file formats, so that images can be loaded into other applications for additional processing. IQWorks provision for image input / output is discussed in more detail in section A.6.

Taking into consideration the key performance criteria outlined above, and weighing up the advantages and limitations of different programming languages, it was decided to write IQWorks using the C# language. C# is intrinsically object-orientated and forms part of the Microsoft .NET Framework. C# source code is prepared as a number of related text files which are compiled into a single 'executable'. However, rather than being native binary code the executable actually consists of 'intermediate language' instructions for the .NET Runtime, a virtual machine environment within which all C# (and other .NET) programs run. Although the .NET Runtime is provided as an integral component of Microsoft Windows, implementations also exist for Linux and Unix (through the Mono platform). Carefully written C# programs are therefore immediately cross-platform, without requiring recompilation. A considerable benefit of .NET languages such as C# is that the inherent provision of the .NET Runtime within Microsoft Windows means that programs written with these can be run without a formal installer being required. Another advantage is that the security around .NET Runtime execution environment is tightly controlled, so that if an errant C# program crashes it should not bring down the entire machine. In addition, all memory management is automatically performed by the Runtime so that memory leaks or violations are extremely rare and less time is required to consider these when writing code. Finally, C# is a 'heavily typed' language, meaning that the types of objects being passed throughout the program are verified by the compiler when the executable is being prepared. For example, if a 32 bit floating point number is expected by a routine then the compiler will verify not just that a number is being passed but that the number is of at least the specified precision. All these extra layers of protection increase the robustness of the final software, which is attractive for a package intended to interact with mission critical servers or medical devices.

One potential issue with C# is that because the intermediate language executable is interpreted by the .NET Runtime, so that there there may be a reduction in number-crunching performance. However, just-in-time native compilation during execution ensures that code runs at native speed where possible, and indeed simple numerical benchmarks indicate this is largely successful. In addition, for very processor intensive tasks, C# code can be written to run

outside of the 'safe' environment or to instead call natively compiled libraries. This flexibility is not possible with similar languages, such as Java.

Comprehensive C# graphical development environments are freely available under both Windows and Linux, making the construction of complex user interfaces straightforward. These also provide tools to aid writing code, including automatic code-completion and semi-automated re-use of existing components, all of which substantially reduce overall development time.

A final advantage of C# is that Varian Medical Systems have opted to write all their software, including numerical algorithms, in C# or other .NET languages. A feature known as '.NET interoperation' allows other C# code to interact at a fundamental level with the Varian executables, without programming interfaces being specifically supplied by the manufacturer, and crucially without installing additional software or interfering with the medical device. IQWorks can therefore be applied to some very advanced research applications, such as the optimisation of CBCT reconstruction algorithms, with very little additional programming.

Microsoft is continually refining and extending the .NET Framework, with the most recent version of C# at the time of writing being 3.5. However, the .NET Runtime provided by default with all Windows XP installations (Service Pack 2 or above) is 2.0. This means that it can be assumed that all Windows PCs, including medical devices, will have full intrinsic support for the .NET Framework 2.0, but not above this. Indeed, the .NET Framework is considered a fundamental Windows component and so some qualified devices are tied to specific versions, but always with 2.0 as the minimum. For this reason, it was decided to write all IQWorks code to be compatible with the .NET Framework version 2.0 (or above), allowing it to be executed on any current Windows computer without requiring the installation of additional components.

Rather than write basic numerical algorithms from scratch, such as those providing the Fast Fourier Transform, it was decided to reuse available validated libraries wherever possible, thus making development more efficient and minimising the opportunity for introducing bugs. IQWorks thus relies on a number of well-established open-source components, as described in section A.4.

Although the core IQWorks system does not require installation, being capable of being run from a memory stick or shared network drive, for full flexibility with certain third party components it is necessary to first register these with the operating system. This is most easily achieved using an automated installer,

which also provides the option for including test images or deploying only specific IQWorks components. Such installations are inappropriate on medical devices, but it was decided that an installer was important both to assist the novice user in getting started and to provide advanced users with additional options regarding the mode of deployment.

It will be clear from the discussion above and in section A.2 that IQWorks provides the user with considerable flexibility over all aspects of its operation, from installing and running the software, through the loading of images, the development, fine-tuning and running of analysis trees, the generation of reports and the storage of results. However, this means that it can be difficult to set the system up to behave exactly as intended. Although the installer has been configured to provide an instance of IQWorks where the default settings for analysis modules are sensible for most modalities, it is not possible for these defaults to be applicable across all phantoms, modalities and applications. It was a deliberate design decision not to hide the complexity from the end user, but to help them work through it. At the heart of IQWorks lies sophisticated image analysis algorithms – all of which are only as reliable as the data and parameters input to them – and it was felt important that users were always aware of the choices they were making. This is in contrast to the approach taken by most commercial packages, which intentionally hide complexity from the user to make them easier to use, but which also has the effect of making it more difficult to identify why results may not be as expected (or indeed, whether any given algorithm is appropriate for the intended application). Considerable effort has been expended in trying to balance the user-friendliness of IQWorks with providing full access to the numerical algorithms, so that they are never reduced to being considered simple 'black-boxes'. The intention is that an experienced scientist will always be involved in developing an analysis tree before it is run in an automated fashion by others. By exposing all possible analysis module configuration options it is hoped IQWorks can help users understand the complexity of the task and be confident in their choice of parameters.

Imaging modalities, clinical applications, test phantoms and analysis methodologies are all continually advancing and it will never be possible for IQWorks to comprehensively accommodate all eventualities. Instead, the goal of the IQWorks release associated with this thesis was to implement a representative cross-section of techniques in the current literature, enabling the basic principles of quantitative image quality evaluation to be applied to any radiotherapy imaging modality. IQWorks can therefore be considered an evolving work in

progress, being a snapshot of the state of the literature at any given time. However, the framework design is intended to ensure that future development and evolution of the system is clear and consistent.

A.4. Software Development

To date, IQWorks has been developed almost entirely by the author using Microsoft Visual Studio Professional, versions 2000–2008. Although it is not a free product, the advanced features provided with the Professional edition of Visual Studio greatly aid in the maintenance of such a large and complex project. Particularly important were sophisticated project management tools, the ability to work with multiple .NET languages alongside C#, and utilities to help build C# code around third party components. However, in keeping with the goal of IQWorks being a freely distributable, open-source package, to encourage contributions from as many people as possible efforts were made at every stage of development to ensure IQWorks can also be developed and compiled using the free 'Express' editions of Visual Studio. At present, the entire IQWorks code base can be successfully compiled without any modification using the free Microsoft Visual Studio 2008 Express.

As discussed in section A.3, care was taken to write code targetting version 2.0 of the .NET Framework so that IQWorks is automatically compatible, without requiring the installation of additional components, with as wide a range of operating system platforms as possible. IQWorks has been successfully tested under Windows 2000, Windows XP (SP2 and SP3), Windows Vista, Windows Server 2003 and Windows 7.

Both to encourage uptake and maximise its potential longevity it is important that the IQWorks program code be clear and accessible. Ideally, a scientist unfamiliar with the project should be able to develop an appreciation of its basic principles through a brief look at the code, then start making changes very rapidly. To achieve this, the program is structured so that each code file corresponds to a logical module within the package, allowing minor yet powerful modifications to be made after gaining familiarity with only a few files. In addition, as far as possible, the program code is written in such a way as to be human readable, using sensible, meaningful names for classes, variables, methods and other objects. Sometimes this leads to rather verbose names, but the auto-complete features of Visual Studio and other editors makes this an efficient way to write and update code, and any additional time spent typing

names is more than saved by the increased readability and thus accessibility. Comments are also included throughout the code, with especially detailed comments being added to help elucidate what might at first appear to be unintuitive algorithmic steps, or where a particular workaround has been implemented to address an unexpected pitfall or bug. Many of these comments can also automatically be used to generate hyperlinked documentation when the program is compiled.

When working on a complex software project it is impractical to rigorously test the entire package each time a minor change is made. Instead, testing is generally focused on the areas of the code which have been modified. However, this can only be reliably performed if one can be confident where the changes have occurred. Some mechanism for recording and tracking changes is therefore crucial. Furthermore, if problems are encountered unexpectedly in the future with a supposedly unaltered part of the package then there must be a mechanism for identifying and investigating all changes made between the last working version and the current point in time. In the worst case, it should be possible to roll back to a previous stable version, and ideally there should be provision for experimenting with new code before committing it to the final project. Mechanisms for tracking and managing changes are therefore crucial, and these were provided using the Subversion version control system.

During the early stages of development the program code was *committed* to a Subversion database, or *repository*. Subsequently, every time a series of changes was made, usually after implementing a new functional unit or resolving a particular bug, the new version would also be committed to the repository, along with a brief textual comment describing the modifications. Subversion does not actually store the full version of the program each time a commit is performed, but rather only the differences between versions. It is therefore possible to very quickly identify exactly what has been altered from one version to the next, or even between any two non-consecutive versions, and to undo changes which are problematic. Version control is even more important when multiple programmers are working on the same code base simultaneously, a common scenario for open-source projects. Subversion includes functionality to ensure that exactly the same files are not being modified concurrently by different users, and also identifies potential conflicts between the code committed by different people. As part of making IQWorks generally available as an open-source project, the code is currently being uploaded to a publicly accessible Subversion repository hosted by the collaboration website Sourceforge.org. Another very

important role of version control systems, which will be discussed in section A.5, is code quality assurance.

IQWorks utilises a number of third party libraries and other modules to provide particular functionality. Where possible, free and open-source components are used which can be fully compiled into the C# executable. In addition, efforts are made to choose truly cross-platform components so that IQWorks remains as portable as possible. If a specific component is closed-source, tied to a particular hardware platform or cannot be widely distributed, then that component is consumed by IQWorks in such a way that the rest of the software can run on all platforms even without this component being present, just not with the functionality provided by it. For example, some proprietary driver libraries are employed to manage the interfacing with display calibration devices (as described in section B.21) and these are closed-source and only function under Microsoft Windows. However, IQWorks will still run successfully without these drivers present, just will not be capable of communicating with those measurement devices.

A list of third party components used by IQWorks is presented in table A.1. One component worth special consideration is 'R', the open-source implementation of the statistical programming environment 'S-Plus'. R is a long established and very efficient numerical analysis package which interfaces well with C#, is cross-platform and can provide computational services over a network. Originally intended for statistical applications it has grown to become a well-respected general analysis package. R is used in IQWorks to provide the vast majority of number-crunching tasks, including Fast Fourier Transforms, interpolation and weighted-polynomial smoothing. A significant advantage of R is that the numerical algorithms are all backed by seminal literature and have been in circulation for such a long time that their implementations can be considered robust and reliable. Furthermore, if a specific analysis routine is unavailable in R but would be more efficiently implemented in R than in C#, then it is possible to write additional plug-in modules to R which can be distributed alongside the C# code.

Once IQWorks has been tested with particular versions of third party components then those versions can be considered 'qualified' with IQWorks. New versions are then only introduced once thorough testing has demonstrated that the functionality they provide is similar, ideally numerically identical, to that previously.

Complementing any software package is documentation to aid the end-

Component	Description
R	Numerical analysis package. Provides core numerical algorithms, including Fast Fourier Transforms, interpolation, etc.
R(D)COM	Interface module linking C# and R.
SourceGrid	Interactive tabular environment for displaying and manipulating data.
iText	Libraries for programmatically generating documents in HTML, RTF (MS Word) or PDF format.
ZedGraph	Charting component.
ClearCanvas	DICOM input/output and networking library.
Microsoft SQL Server Express	Comprehensive database server.
OpenJPEG	Codec for decompressing JPEG 2000 images.
Independent JPEG Group Toolkit	Library for decoding JPEG images compressed with lossless and advanced lossy algorithms not covered by the .NET framework.
CVIPTools	Library providing advanced edge detection and other image processing algorithms.
SourceGrid	Provides spreadsheet-like interactive grid for displaying and editing data.
AVIFile	Enables saving of multiple images as a cine movie loop.
Icons - Silk and Fugue	Icons used in the graphical user interface.
InnoSetup	Generates all in one installer package.
Instrument Drivers	Interfaces for communicating with the X-Rite Eye-One Display and IBA LX series of photometers.

Table A.1. – Third party components utilised by IQWorks.

user leverage the software to its full potential, or to help the programmer become quickly orientated with the code so that they can begin contributing to the project. This is an area currently being addressed in IQWorks, with on-line documentation being built to complement training courses and scientific meetings.

A.5. Validation and Code QA

For any numerical analysis package it is important that the results generated under well-defined conditions are accurate, predictable, stable over time and reproducible from one software release to the next. However, this is absolutely crucial for a package such as IQWorks which may underpin an imaging QA

programme and where measurements may lead to time-consuming and costly maintenance of equipment required for patient treatment, potentially controversial technical comparisons between units manufactured by different vendors, and in extreme cases, to a decision to alter the management of a patient's disease. As long as it is utilised by trained scientists or professionals as a QA tool to inform their decision making, then IQWorks is not technically classed as a medical device, yet the reliability of its results is just as important if it is to provide an overall benefit and not be a hindrance, particularly as it becomes relied upon to support complex, objective analysis schemes. Continual validation of IQWorks for particular applications is therefore key.

Validation involves two considerations: ensuring that the measurements produced under standard conditions are accurate, and that these results do not change over time or between software revisions. Although related, these are actually very different. Accurate performance of individual analysis algorithms, and the way in which these work together when linked in an analysis tree, is essentially verifying that the programmatic implementation of the algorithms and the IQWorks framework produces results in line with theory. However, there may be uncertainty or bias in the results due to factors such as rounding effects inherent to floating point arithmetic, assumptions or default choices based on current thinking in the literature at the time when an algorithm was implemented, or even simply exactly how the graphical representation of a region of interest corresponds to the extraction of a matrix of numerical data (e.g. are pixels under a line denoting the edge of an ROI considered inside the region or outside?). On the other hand, reproducibility over time requires a stable implementation of the algorithms and ensures that trees developed and relied upon for a given application are still valid when a new software version is released. Although arguably of more importance to the end-user, reproducibility can be difficult to achieve. If best-practice in the literature results in new default values being considered more appropriate, then there can be strong motivation to change these to accommodate. Alternatively, more efficient numerical implementations may result in different rounding effects which can accumulate to give slightly, yet significantly, different results. Lastly, a fix to an identified bug may also cause a difference in results, even under conditions where the bug would not previously have had a noticeable impact. Because of the importance of stability and the ability to collect reliable time-trend data, it was decided in IQWorks to ensure, unless absolutely necessary, that reproducibility be maintained between releases, even at the expense of

making the software slightly more difficult to use. For example, if a new option becomes available from an existing drop-down list, or an enhancement to an algorithm presents the user with new choices, then IQWorks attempts to mimic the previous version of the software by ensuring that when existing trees are loaded all selections are the same as before and any enhancements or new options are by default turned off. Furthermore, where a bug fix corrects anything other than a major algorithmic error there will usually be an option provided to turn off the correction so that the user can assess the implications of it for their particular application. However, bug fixes are by default always turned on.

Each of the analysis modules described in the following appendix was tested to ensure correct behaviour, and a range of analysis trees was constructed to verify the correct operating of IQWorks input / output, reporting and storage functionality, along with the interaction between the different modules. Testing was performed using a combination of real phantom images, acquired using different modalities, and artificial images generated to test particular functionality. Throughout the IQWorks development process, a reference set of test images was collected and a standard set of analysis trees constructed. Whenever major changes were made to the code, and prior to each software release, the standard trees would be run against the reference images and the results compared with those attained previously. As IQWorks becomes more complex this testing is progressively more onerous, yet this system-wide testing is even more important if serious bugs are to be identified and multi-user contributions are to be rigorously tested as the open-source initiative gains momentum. It is intended that an automatic testing framework be developed to extend systemic testing so that large numbers of trees can automatically be run against large numbers of reference images, thus assuring the quality of the package for standard applications.

Actual tests for each analysis module are described in more detail alongside their implementations in appendix B, and the analysis trees are discussed with the modality examples in chapters 4 to 5. However, only the validation of the more complex algorithms is considered in depth, rather than basic statistical or arithmetic calculations which are straightforward to verify.

Determining whether the results generated by a particular analysis module are 'correct' is not always straightforward, due both to numerical implementation issues and design decisions, as described above. Where the results are well-defined by underlying theory, the performance of a routine is deemed

acceptable if it agrees well with theory to a precision at least equivalent to, and ideally greater than, that required for the intended clinical application. For more complex analyses, or those underpinned by more ambiguous theory, results are compared with those yielded by packages based on other authors' work, some of which are described in section A.11. Clearly, other packages are themselves subject to validation and their own implementation nuances. Greater reliance was given on results generated by the DIMOND3 tools and the ImPACT+ package because these were developed by well-established groups, are backed by significant literature and have been received positively by the community.

Following the validation process, a user can therefore be confident that IQWorks incorporates robust implementations of its image evaluation algorithms and generates results comparable with those of other researchers. However, this does not mean that IQWorks is automatically suitable for any given clinical application, with the onus being on the end-user to perform whatever testing as they deem sufficient to verify IQWorks meets their particular needs. Although the processes for basic image quality evaluation for many modalities are now relatively mature, so that only a small number of phantoms might be utilised as part of a routine QA programme, there will always be local peculiarities to take into consideration. A question sometimes raised is the issue of who is liable should IQWorks produce incorrect results which then results in harm to a patient, albeit indirectly because IQWorks is intended only to assess the technical performance of equipment not to actually make decisions regarding its suitability for use. It is recommended by the author that, as it develops as an open-source project, IQWorks be considered a peer-reviewed implementation of standard image evaluation algorithms which can be applied if a healthcare professional understands the technical details of the package and validates it for their application. In the same way that a healthcare professional might decide to introduce into their practice a new technique from the literature, IQWorks should only be used after extensive testing, assessment of suitability for purpose and investigation of its limitations. Because the full source-code is available, along with test images and example trees, the end-user has all the information required to determine whether or not the package meets their needs. In many ways, this is a better situation than with closed-source commercial implementations, where the details and limitations of algorithms may never be fully understood and yet the same requirement of user-validation applies.

Even when programming very carefully bugs will inevitably appear in soft-

ware, including rounding errors, inefficient or erroneous numerical implementations of algorithms, or issues over memory allocation (even though these are generally dealt with well by the .NET Runtime). Visual Studio provides comprehensive debugging tools which allows code to be stepped through line by line whilst monitoring the state of key variables. In addition, the ANTS Profiler was used to identify redundant code and loops which were being executed more times than necessary, which can be a significant issue with object-orientated software where the assumption is that each developed module does not need to be concerned with the internal operation of others. Memory profiling by ANTS was also invaluable in catching large blocks of memory which were not being released when images were no longer active, and which thus were limiting how many images IQWorks was able to process simultaneously.

A very powerful method of ensuring that code performs as intended, and continues to do so even after changes are made, is 'unit testing'. This involves writing specific, dedicated routines to test new code functionality at the same time as the new code is being written. Visual Studio and other development environments include tools to automatically perform these tests and flag up any identified problems. In the extreme case, the tests may be written in advance of the code, with the philosophy being that all functional requirements should be well defined and that the purpose of the code is to meet these. Developing unit tests adds a significant time burden to the writing of software and is not a practice normally followed in scientific programming. However, it is a methodology almost universally embraced in the professional software engineering world, with it being felt that the time gained from having to tackle less bugs is more than justified by the time required to develop the unit tests. Only basic unit tests were incorporated into the IQWorks code base during initial phases of development, but these are to be extended as the project takes on new developers.

Subversion was introduced in section A.4 as a means of tracking and managing code changes. In any complex project, detailed revision tracking enables an experienced programmer to identify exactly which lines of code have been modified from one revision to the next and thus determine the testing regime required to repeat validation. Furthermore, if many users are simultaneously making changes to a public repository, then the revision control tools enables the implications of all these changes to be assessed together before it is decided whether or not to accept them permanently. As IQWorks develops, it is intended that a small team of experts be formed to safeguard the reference code

implementation that is the basis of the official releases. Any programmer will be permitted to make changes, then members of this team will peer-review the new code to verify the contributions are of sufficient quality and merit to be included in the official release. Whenever a new version is released, the physicist responsible for its implementation in the clinical environment can access the associated *diff* file describing all the differences between the current and previous versions. Alongside the official release notes, this will help guide which tests are required before the physicist can be confident that the new version continues to meet their needs.

A.6. Image Input / Output

Any image evaluation task begins with loading the images concerned into the analysis package. IQWorks is capable of interpreting image files stored on any media accessible by the operating system, including locally installed hard disks, CD-ROMs, DVDs, memory sticks and network shared drives. Where images are required from a networked database store, such as a PACS, these can be accessed via any third party tool which queries the database and outputs files to disk. Although it would be relatively straightforward to incorporate such networking functionality into IQWorks itself it was felt that because this is already adequately provided by a number of readily available open- and closed-source packages, it was more important to concentrate in the initial stages on the reliable interpretation of disk-based file formats. However, direct querying of databases has the potential for making some applications more efficient so has been tabled as a worthwhile future evolution. For example, if IQWorks could automatically determine the locations of images stored in the Varian ARIA radiotherapy information system by querying the ARIA database, then these could be accessed directly instead of there being a requirement to manually use the ARIA applications to first export copies of the images to another location.

Numerous file formats exist for storing medical images, with the format chosen in a particular instance depending upon the imaging modality, computer hardware and operating system, the nature of the image data being stored (both pixel data and additional information), programmer / vendor preference and other historical reasons. However, an increasing number of modalities are moving to the DICOM 3.0 format for the files, and almost all modalities enable images to be exported in this format if it not used natively.

DICOM (Digital Imaging and COmmunications in Medicine) 3.0 is a standardised protocol for the storage and exchange of imaging information. It is managed jointly by the American College of Radiology (ACR) and the National Electrical Manufacturers' Association (NEMA). Although the standard is referred to as 'DICOM 3.0' this is misleading, with the number '3.0' indicating that this is the third iteration of a fundamental standard, with the DICOM 1.0 effectively having been the Interfile format still widely used in Nuclear Medicine, and DICOM 2.0 being the ACR-NEMA 2.0 format. DICOM 3.0 is actually a continually evolving set of protocols which is formally updated and released every six months, sometimes with substantial revisions and not always to be completely compatible with older versions. Intended to be applicable to all current and future imaging modalities, DICOM 3.0 is an extremely wide-ranging and complex standard, not limited to just the definition of an image file format, but rather covering all aspects of the information transfer process, including network protocols, physical aspects of storage media, binary transfer syntaxes and display device calibration.

In addition to image data, DICOM files can represent other types of information associated with images, such as procedure worklists, radiotherapy structure sets, radiotherapy dose matrices and radiotherapy treatment plans. Each DICOM file contains only one type of information, although there may be multiple instances of this, such as a stack of MR slices in a volumetric image. Along with the core information item there is also always *metadata* – information placing the item in context – such as an identifier describing what the item represents, the date and time of acquisition, the manufacturer's name and the user identifier of the operator.

Even when just considering a simple image file, there are optional data elements which can be included or omitted depending upon user requirements, with the choices made then resulting in a different set of mandatory and optional components. In addition, because the specification for a particular modality is developed by a group of experts working in that modality, the options available vary between modalities and there can be conflicts or ambiguities when an imaging device may acquire images belonging to more than one traditional modality: for example, CT images may be considered a traditional CT scan, a volumetric cardiac image set or a scan intended for use in radiotherapy. Furthermore, there is provision for manufacturers to include their own proprietary information, a facility often used when new features are introduced which are not currently supported by the main standard.

For all these reasons it is extremely difficult to develop a comprehensive and robust implementation of the DICOM standard capable of reading and writing all DICOM files in existence. Even when a piece of equipment is designated as 'DICOM Compliant' or 'DICOM Compatible' there is considerable flexibility over the nature of the implementation. Indeed, sometimes assumptions are made the files are expected to be read back by equipment from the same manufacturer, leading to some of the information stored in the file actually being wrong. However, these may not always be obvious or clear to the end user and third party applications such as IQWorks. This is especially true in radiotherapy imaging where key information may not be encoded properly, such as the location of the central axis in a projection image, simply because the applications intended to interpret these images can generate it by other means. To alleviate compatibility issues, the DICOM standard specifies that a 'DICOM Conformance Statement' should be provided by a vendor which describes exactly which parts of the standard have been implemented, and how. However, these statements tend to be long, complicated documents and the responsibility lies with the user to verify the standard is implemented as expected. In particular, because conformance statements tend only to discuss implemented functionality it can be unclear which key functionality may actually be missing. Generally, conformance statements are of greatest utility when trying to identify why two supposedly compatible systems are not communicating properly.

A pragmatic approach was therefore taken when writing IQWorks' DICOM import functionality. By default, image information is loaded with complete fidelity to what is contained within the image file. If fundamental problems are later identified by the user then the import code can be modified to recognise the particular type of image and apply a correction during the import process. For example, Varian Ximatron Simulator-CT images are often stored with the origin of coordinates in the wrong location. Hard-coded into the import module is therefore a clause to identify Ximatron images and force the origin to the centre of the field of view. Another common scenario is when the information stored with the image data, although correct, is not what the end user was expecting. For example, if images of phantoms are acquired but not actually viewed on the modality in question before export, then the window settings stored with the image data may not display the image data so that particular features are visible to the user. In attempting to faithfully interpret the data stored in the file, IQWorks will correctly apply the stored window settings, presenting the user with an image which does not at all appear as expected. A

further example is when the user is expecting the origin of an image to be at the centre of the field of view, but it is actually elsewhere. Although the system is behaving as intended, this can be an intense source of frustration. For these circumstances, global override options are provided which enable a 'correction' to be applied to all loaded images in preference to what is actually contained within the headers. By default, all such overrides are turned off, but IQWorks preserves a user's preferences from one session to the next.

Core DICOM handling is performed using the Clear Canvas libraries. These were chosen because they provide a .NET implementation of many aspects of the DICOM standard and form the basis of a well-established and robust open-source PACS system. A particular advantage of the Clear Canvas libraries is that they can automatically and reliably decompress the whole range of DICOM compressed image formats. In addition, a comprehensive implementation of DICOM network services will enable this functionality to be incorporated into IQWorks in the future. Because Clear Canvas is supported and maintained by an active community it is expected that it will keep abreast of new developments and regularly be updated to include functionality expressed in newer versions of the DICOM standard. Therefore, in utilising the Clear Canvas libraries to handle DICOM input / output it is anticipated that the broadest possible DICOM support will be achieved into the future, without modifications being required to IQWorks itself.

When interpreting an image file, IQWorks first uses Clear Canvas to extract all metadata information stored in the file. This is stored and indexed by item *tag* name so that the user can interrogate the metadata whilst viewing the image. Before actually decoding the image data, IQWorks checks for the presence of key tags in the metadata. These include the modality type, dimensions of the pixel matrix, the number of images stored in a file, scaling information to convert the stored pixel values to actual ones, pixel size, location of the origin and image orientation with respect to the patient's internal coordinate system. If key information is missing then IQWorks makes certain assumptions. For example, if pixel size or scaling information are not present then these are assumed to be unity. If frame of reference information is missing then the image is assumed to be an axial slice acquired with a patient lying head-first supine, with the origin of coordinates being set as the centre of the field of view. More information about coordinate systems and frames of reference is provided in section A.7. Metadata tags are also checked for identifiers of problematic image sources where 'tweaking' of the image data may be required, as described above. Once

the context of the image has been determined the pixel data themselves are decoded and adjusted accordingly before being stored in memory.

If a DICOM file contains multiple images as either a time-series or an inherently 3 or 4-dimensional dataset, then these are loaded and stored together as a multi-planar image. IQWorks also allows the user to select multiple image files simultaneously, giving the user the choice of interpreting them as a series of images to be considered individually within the package, or to all be aggregated into a single multi-planar image. Multi-planar images are described in section A.7.

A common scenario when working with many images is for the end-user to be faced with a series of directories containing thousands of cryptically labelled files. This may occur when images have been received over a network and dumped to disk, or when interrogating the directory tree of a PACS system. IQWorks provides an image preview facility which displays a thumbnail of any file when it is clicked on in the open-file dialogue box. Although this is a useful aid to selecting the correct image when one knows the location but is unsure of the filename it is not practical for sifting through a very large number of images. Instead, a 'DICOM Browser' tool has been written which scans a given directory and constructs a catalogue of the files within it. A screenshot of the DICOM Browser is shown in figure A.3, which illustrates the results of scanning a directory of standard test images acquired using a range of modalities.

After the user selects a root directory for scanning, the DICOM Browser constructs a list of all files in this directory, also working recursively through all sub-directories if required. For each of the files in the list, the metadata is interpreted and items of information important for constructing the catalogue are extracted, such as the unique identifiers, time stamps and names of the image study, series and individual instance. These are stored in memory alongside the filename. At this point IQWorks does not attempt to load any image data because it is likely that the majority of files in the directory will not require to be processed, so that loading the image information would be an unnecessary burden on both time and memory. If a file cannot be successfully interpreted, for example because it is not a DICOM file or has been corrupted, then the browser simply ignores it. Once all files have been examined the unique identifiers of the studies, series and other objects are cross-referenced to build the catalogue, which is presented as shown in the figure.

Initially, a list of studies is presented. When the user clicks on a study, the series belonging to that study are shown, then when a series is selected

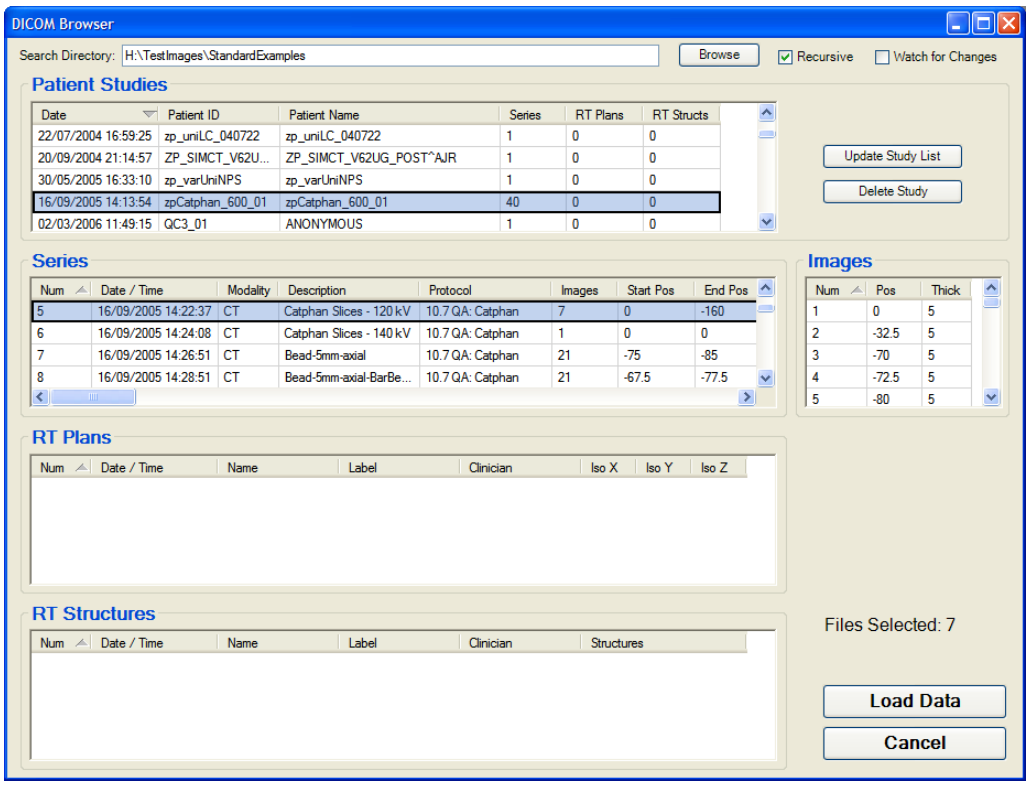


Figure A.3. – Demonstration of the DICOM browser illustrating the results of scanning a tree of DICOM files then selecting a series of Catphan 600 images.

the associated images, radiotherapy plans and radiotherapy structure sets associated with that series are listed. The user can select an individual image, multiple-images, or individual or multiple series or studies. For example, in figure A.3 the whole of series 5 of the zpCatphan_600_01 study has been selected. Both individual images and large numbers of images can therefore be selected and loaded very quickly. Each time the selection changes, the number of underlying images which would be loaded is displayed so that the user has an indication of the memory and time implications of loading these. Radiotherapy plans and structure sets are included in the browser listing because these are used by the DRR analysis module, which is covered in section B.20. Once the selection is complete and the user initiates loading of the images, the browser tool passes the list of selected files to exactly the same routines as for loading images manually.

Image repositories typically contain many thousands of images so that performing a full scan each time the DICOM Browser loads can be very time-consuming, yet it is important that images added since the previous scan are detected and added to the list. For example, the repository of daily EPI check

images in Oxford consists of over 3000 images and occupies more than 4.5 Gb of space. By default, IQWorks maintains a cache of the DICOM Browser catalogue which is maintained between sessions. All information necessary for building the catalogue, plus file sizes, time stamps and a note of the last time the file was fully scanned, is recorded in the cache. Each time a directory is scanned, the browser checks whether a file is in the cache. If the file is present, it verifies that the last scan time is not outside a pre-defined 'expiry' period, then compares the file size and modification time stamp against those in the cache. If the file is not present, the previous scan time has expired, or the file size or modification time stamp have changed, then the file is scanned again. Otherwise, the file is not scanned and the information in the cache is used instead. This caching facility considerably speeds up population of the DICOM browser, with the limiting factor being the time taken to construct the list of files under the specified directory. Generally, this is of the order of only a few seconds, even for very large, nested directory trees.

Another feature which aims to make interrogation of large image repositories more efficient is the ability to monitor an entire directory tree for changes and update the browser automatically when one is detected. This is performed using an operating system hook built-in to the .NET framework and is extremely powerful because deep directory structures can be monitored with a negligible time penalty. This functionality is particularly useful when waiting for an image to arrive from a QA device or PACS system.

An option is also provided in the browser for deleting whole studies or series from the directory tree, a task which can be difficult when the files on disk do not have meaningful names.

Originally, the DICOM browser was developed with a range of radiotherapy applications in mind and the information which is displayed was thus tailored to these. However, it is acknowledged that other metadata within the DICOM files may be more useful for other applications, or for modalities not commonly encountered in radiotherapy, and efforts are underway to extend the tool to improve its overall flexibility. In particular, it would be advantageous to display more detailed information regarding the individual images themselves.

In addition to DICOM files, IQWorks can successfully interpret a range of bitmap files, including those compressed according the JPEG 2000 and JPEG-LS lossless standards, Interfile, ACR-NEMA 2.0 and Varian native image formats. Furthermore, ASCII files and raw binary files can also be loaded, although the user must provide basic information about the structure of these, such as

encodings and field delimiters.

Regardless of how flexible IQWorks is at processing and evaluating images there will always be occasions when it is necessary to export an image for import into another package. All IQWorks images, including those which are generated by tools within the package or specific analysis modules, can be exported as bitmaps, ASCII files or raw binary files. Sometimes this functionality can be useful simply for converting DICOM or other images into a more convenient file format for inclusion in presentations, etc. Furthermore, images which have an analysis tree attached can also be saved as bitmaps which include the graphical overlays belonging to the analysis modules within the tree.

A.7. Image Handling

Pixel data are central to all IQWorks operations and are represented as an array of 2D matrices of IEEE single precision floating point values (i.e. an array of C# float[,] types). For example, a simple 2D image corresponds to a single float[,] matrix of dimensions the same as the image, whereas a volumetric image is represented by an array of float[,] matrices, one for each of the planes in the 3D image. Except for processor intensive operations, float[,] types are never addressed directly in the code, but rather are encapsulated by IQWorks 'FloatArray2D objects'. These incorporate boundary checking to prevent reading or writing beyond the limits of the matrix, as well as additional routines to perform common statistical calculations.

Images comprising more than one FloatArray2D object are referred to in IQWorks as *multi-layer images*. These are intended to represent multiple 2D planes which are related to each other and will be considered simultaneously by any particular analysis module. For example, a stack of Catphan 504 CT slices may be processed as a multi-layer image by the slice-sensitivity analysis module, or a sequential series of flood EPIs may be used as input to the NPS algorithm, where regions of interest across more than one image will contribute to the ensemble average. Depending upon the nature of the analysis module, when a multi-layer image is loaded the user is given the option of choosing to process either a specifically numbered layer, the layer which is currently visible, the first layer in the stack, or all layers present. In this way, different analysis modules in the same tree can act on different layers of the same multi-layer image.

Implicit to the representation of an individual pixel by a single floating point

number is that IQWorks can only inherently handle greyscale images. However, depending upon the application, colour images can either be loaded as a multi-layer image (in which each layer corresponds to one of the red, green and blue channels, for example) or can be collapsed into the equivalent greyscale (or luminance) image. This approach lends itself better to standard techniques for processing colour images, where the individual components are usually considered separately. It is important to note that the planes of a multi-layer image need not have been acquired at the same time, or even by the same imaging device, with the only criteria being that the different planes have the same dimensions and pixel size. Routines are provided both to combine individual images into multi-layer images, and to separate multi-layer images into their constituent planes.

Multiple images can be loaded into the IQWorks environment simultaneously so that the same analysis tree can be run against each sequentially. Each individual image can itself be a multi-layer image. For example, an experiment might involve calculating the NPS for EPIs acquired on different linacs for a set of acquisition settings. All the images can be loaded as multi-layer images, then the appropriate analysis tree automatically run against each of these, one after the other.

Internally, there is separation between the routines for managing and processing image data, and those for display images. In line with the object-orientated paradigm discussed in section A.3, this approach enables modifications and refinements to be made to one without them affecting the others. It will also facilitate the processing of images automatically by a background service, perhaps across the Internet, which is an intended future evolution.

Images are displayed in IQWorks using a specially written 'ImageDisplayForm' as shown in figure A.4, which illustrates a volumetric cone-beam CT image of the Catphan 504 phantom acquired using the Varian OBI. The blue cross-hairs indicates the origin of the physical, real-world coordinate system and the file name of the image is shown in the title bar. If the image is a multi-layer image (as in the figure) then a slider tool is automatically presented which allows the user to choose which plane to view. This can also be achieved by manually entering the index number of the desired plane. It can sometimes be useful to visualise a stack of images as a movie loop, and this functionality is provided by the arrow button next to the slider tool. Furthermore, IQWorks allows saving of multi-layer images as AVI movie files for inclusion in presentations.

Any particular hardware platform is only capable of displaying a limited range of grey levels, with most radiotherapy display devices being limited to 256 unique values. On the right hand side of the ImageDisplayForm is a histogram showing the distribution of pixel values in the active plane, with the black arrows indicating the range of pixel values currently being mapped onto the grey levels available to the hardware display device. The pixel values in the indicated range are mapped with equal spacing across the entire range of available display levels. This *window* onto the image data can be interactively adjusted by using the mouse to move the bars, and can also be set manually through entering its central value (the *window level*) and range (the *window width*). Pixels with values below the lower limit of the window are set to zero (displayed black) and those above the upper limit are set to the maximum available display level (displayed white).

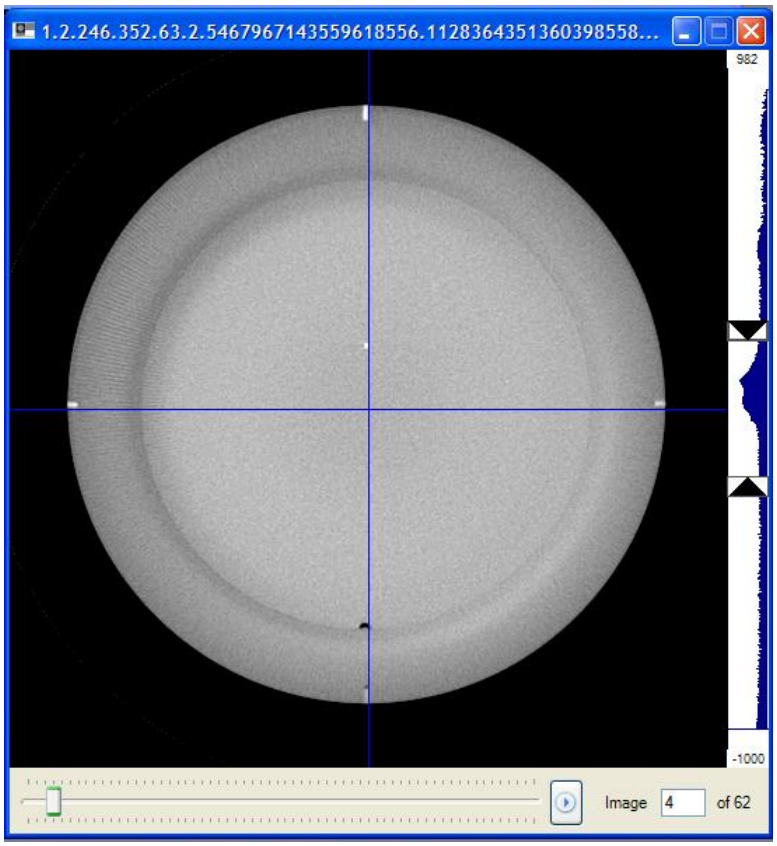


Figure A.4. – IQWorks 'ImageDisplayForm' showing a Varian OBI cone-beam CT scan of the Catphan 504 phantom. The slice containing the Catphan impulse bead is currently active.

To provide the user with flexibility when constructing analysis trees, it is possible to fine-tune all aspects of the geometrical placement of analysis modules.

Internally, IQWorks uses two coordinate systems:

- the pixel coordinate system, which has its origin at the top-left most pixel in the image and describes the displacement of a point as the number of pixels in each matrix direction;
- the physical coordinate system, which has its origin at a customisable pixel location, takes into account real-world pixel size, and represents the displacement from the origin in millimetres in each direction.

Both coordinate systems increase from left to right and top to bottom in the plane of the image, with X representing the horizontal component of vectors and Y the vertical component. Depending upon the application, the user can choose to work in one coordinate system or another, and IQWorks automatically and transparently provides the transformation between them. For example, the location of the impulse bead in the Catphan 504 is positioned 2.0 cm above the centre of the phantom and centred laterally. Therefore, when placing a region of interest to determine the point spread function, it would be logical to centre it at $(0, -20)$ in physical coordinates, ensuring that the impulse bead is always at the centre of the ROI regardless of pixel size or how the reconstructed field of view relates to the physical phantom. Alternatively, an object placed on the surface of an EPI or CR plate may always be a particular number of pixels from the edge regardless of the location of the detector with respect to the x-ray source or central axis, in which case the pixel coordinate system would be more appropriate.

Once an image has been loaded it is possible to obtain more information about it from the *image properties information box*, as illustrated in figure A.5. The file name of the image is presented in the title bar, then below this the name and ID of the patient, followed by the date and time when the image was acquired or generated. If the 'Geometry' tab is selected, basic information is shown about the dimensions of the image, the pixel size in each matrix direction (in mm) and the pixel which represents the origin of coordinates of the physical coordinate system. Both pixel size and origin coordinates can be modified by typing new values into the boxes. Alternatively, the user can set a new origin position by clicking directly on the required location in the image. All metadata extracted from the headers of the image file can be viewed in the 'Information' tab and for DICOM images there is usually a considerable amount of information present.

Objects are placed relative to a specified *anchor* position. By default this is the origin of the chosen coordinate system. However, the anchor can also be

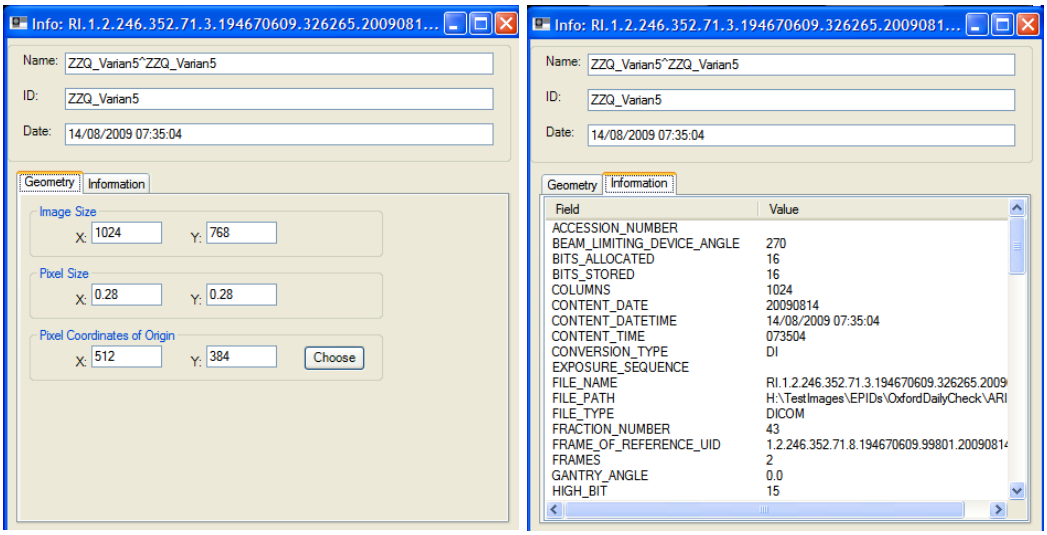


Figure A.5. – Image Properties box, showing geometry and metadata information for a Varian aS1000 EPI of the QEPI1 phantom.

set to be any point output by an analysis module earlier in the tree, such as the minimum or maximum point within an ROI, or the centre of mass of a detected phantom edge. In this way, the placement of objects can automatically be adjusted depending upon where the phantom or test details lie within the field of view. Many analysis modules output a generic *Processed Point*, the nature of which depends upon parameters intput to the given module. For example, the processed point of an edge detector can be the centre of mass, centre of extremes or even just the centre of the search ROI. By setting the anchor of a set of analysis modules to be the processed point of the edge detector, the actual reference point used by all dependent modules can be adjusted simply by changing the definition of the processed point. This is useful when a particular image proves difficult to process and ensures a successful result whilst requiring the minimum of user alterations to the analysis tree.

Aside from the passing of inputs and outputs, all analysis modules are designed to function independently of each other, with each having their own parameters governing how the processing of an image should proceed. However, sometimes it can be useful to condition the underlying pixel data before any analysis is performed, with the result affecting all analysis modules and trees run against that image. For example, if a test phantom has accidentally been positioned upside down it may be desirable to first 'flip' the image vertically to correct this error, or if an image is suffering from a cluster of bad pixels then a median convolution filter might be run to mitigate this. IQWorks

therefore provides a set of immediately applied operations for fundamentally modifying the underlying raw pixel data. These include horizontal and vertical flipping, and different sizes of mean and median convolution filters.

When working with multiple images it can sometimes be convenient to group these into collections for further study or to act as reference datasets against which future analyses are compared. For instance, locating the appropriate image plane in a sizable volumetric stack for a particular type of analysis can be relatively tedious, especially when the stack itself is loaded from a larger directory of DICOM files. Once the image has been located it would be useful to transfer it to a separate location for easier subsequent access. IQWorks provides a facility for copying currently loaded images into a new location so that such collections can easily be constructed.

A.8. Signal Calibrators

Mapping of raw pixel data to linearised 'calibrated' values is performed using *signal calibrators*, which essentially reverse the characteristic curve of any given imaging system. IQWorks provides a number of signal calibrators which can be fine-tuned to model the response curves of the most commonly encountered imaging modalities. Following object-orientated principles, all signal calibrators inherit from the same abstract base class. Therefore, regardless of the parameters or implementation details of any specific calibrator, all calibrators can be managed and applied through a common user interface. In addition, this approach makes developing new signal calibrators for modalities not accommodated by those already present very straightforward, ensuring that only testing of the performance of the new calibrator is required, not its interaction with other parts of the system.

Signal calibrators do not actually modify the underlying pixel data, but rather apply a customisable curve to convert the signal at any particular pixel location into its calibrated value. Analysis modules do not have direct access to raw pixel data, only to the values output by signal calibrators. Applying a calibration curve every time a pixel was addressed would incur a considerable performance penalty. Therefore, the first time any pixel value is requested through a calibrator, the calibrator calculates the values for all pixels in the image and stores a working *calibrated image* in memory. Although this effectively doubles the memory requirements for any individual image this is more than compensated for by a significant performance gain. Signal calibrators are

programmed to identify the special case when the calibrated image is identical to the raw image, in which case a calibrated image is not generated and the raw pixel values are instead passed through to the calling routines unchanged.

To date, three signal calibrators have been implemented in IQWorks, each of which converts the pixel value PV into a calibrated value CV. Although the purpose of these is to undo the characteristic curve, their parameters are described from the perspective of the characteristic curves themselves, because it is the characteristic curves which would be measured by the end user.

Many modalities are inherently linear, including EPI, CT and DR. These are best represented by the *linear signal calibrator*

$$PV = m \cdot CV + c \tag{A.1}$$

where m and c are the constants, being the gradient and DC offset respectively.

CR plates tend to have a logarithmic response and are accommodated by the *logarithmic signal calibrator*

$$PV = m \cdot \ln(CV) + c \tag{A.2}$$

whilst these and other detectors may sometimes be better described by the *power law signal calibrator*

$$PV = m \cdot CV^b + c \tag{A.3}$$

where b is also a constant and is usually ~0.5.

If a signal calibrator is inappropriately chosen so that the calibration curve would result in singularities or invalid computations (such as applying equation A.2 to the negative values of a CT image) then IQWorks sets the calibrated values of the affected pixels to zero.

Parameters for all signal calibrators are modified through the same user interface dialogue, as illustrated in figure A.6 for the logarithmic signal calibrator. This also enables a unit to be defined for the calibrated values (such as Gy if the calibration is being performed in terms of air kerma, for example) although this unit is only used when reporting results and doesn't influence the actual processing by analysis modules. Once a signal calibrator has been configured its parameters can be saved to disk for quick reloading the next time it is required.

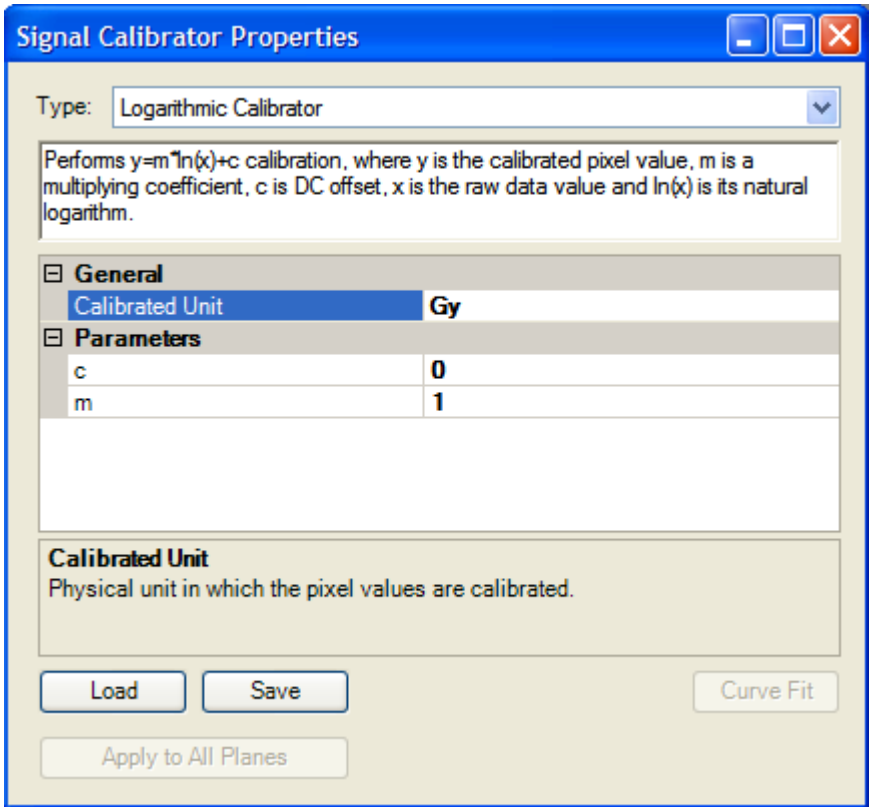


Figure A.6. – Example dialogue box for adjusting the parameters of the logarithmic signal calibrator, where the calibrated pixel values would be represented in Gy.

A.9. Results Reporting and Testing Framework

When processed, all IQWorks analysis modules generate a short summary of their key results which is displayed in the 'Results' pane of the Properties Dialogue box, as illustrated in figure A.2. In addition, all available outputs from a module are listed as read-only results after all user-editable parameters in the Properties Dialogue Box. However, additional reporting functionality is necessary when interpreting the results of running a complex analysis tree, or to meet the reporting requirements of a routine QA programme.

Incorporated into IQWorks is a 'Detailed Report' module which is designed to gather image information and desired analysis module outputs and present these as a structured report in a standard format. Any analysis module output can be included in the report – numerical, textual or graphical – and reports can be generated as rich-text format (RTF – Microsoft Word), Portable Document Format (PDF – Adobe Acrobat) and using HyperText Markup Language (HTML – Web Browser). Programmatically, all three document formats are generated

using almost exactly the same function calls, with the underlying iText library being utilised to manage the document layout.

Reports are constructed using the *Report Generator* module, which is typically added to the end of an analysis tree. An example of the Report Generator being applied to develop a report describing the results of the analysis of the QEPI1 portal imaging phantom as part of the Oxford Cancer Centre daily linac checks is shown in figure A.7. The user begins by entering a title for the report, along with any descriptive comments that should appear as text below the title. If required, a copy of the image can also be included and any useful items of metadata from the image file. Sometimes, the technical 'official' names of metadata elements can be cryptic and unhelpful, so for every item which is included the user can enter a 'friendly name' which will be printed alongside the contents of the item in the report.

All available analysis module outputs, ordered in the same way as the modules they belong to are included in the analysis tree, are displayed at the top-left of the Report Generator. Using the mouse, a tick mark is placed next to each output to be included in the report, and the formatting of each numerical result can be tailored individually in terms of its number of decimal places and whether scientific notation should be used.

As well as simply being displayed, numerical and textual results can have tests performed on them to give 'pass' or 'fail' results. This can be useful for quick quality control checks where it is desirable to verify whether an image quality metric is within certain parameters. Available tests include checking whether a result is above or below a reference value, within specified limits, or exactly matches a particular value. Test criteria can be customised to automatically change depending upon information contained within the image metadata. For example, in figure A.7, different tolerances on spatial resolution (in terms of the spatial frequency at which the MTF reduces to 50%, the f_{50}) are defined depending upon whether the linac photon energy is 6 or 15 MV.

Figure A.8 shows the PDF report prepared by the Report Generator in figure A.7. Following the report title and any comments entered by the user (none in this example) is a summary of the number of 'pass' and 'fail' test results which were evaluated for the outputs included in the report. When multiple images are processed sequentially this summary includes the test results from all the images being considered. Summarising the state of the test results at the start of the report enables the user to immediately identify whether there were problems with any of the tests. For a rapid QC check, the user might proceed

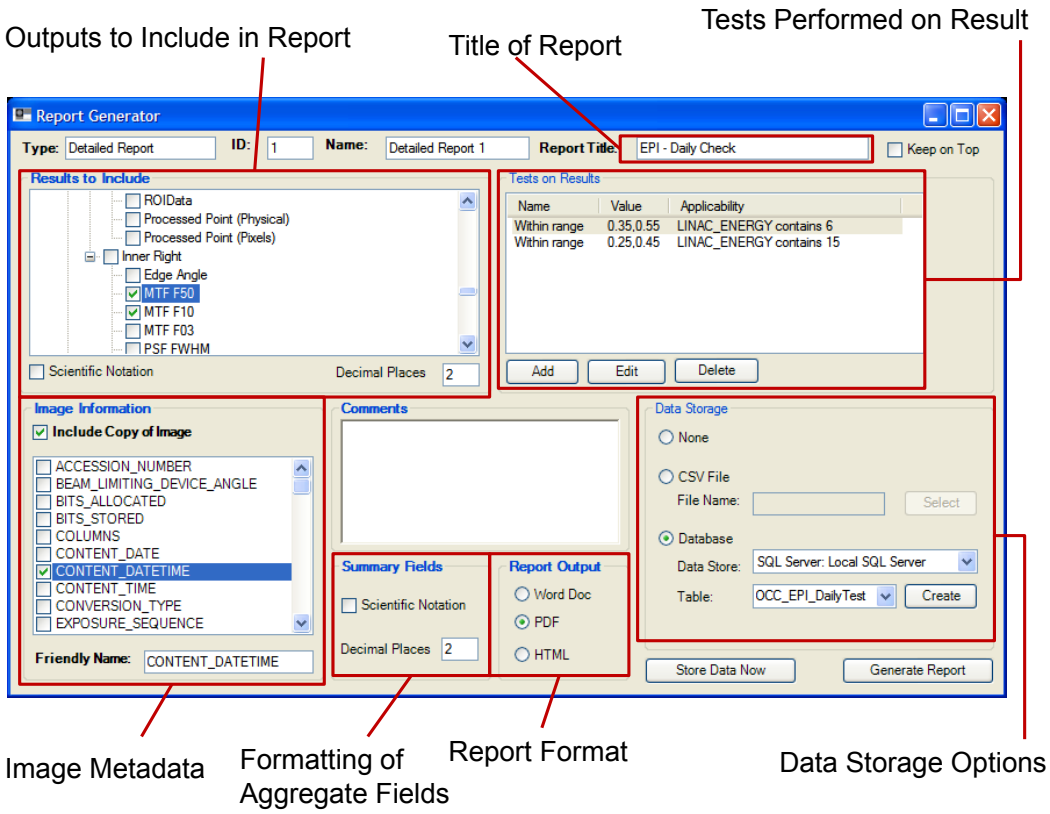


Figure A.7. – Report Generator dialogue box, demonstrating the construction of a report for the analysis of the QEPI1 portal imaging phantom as part of the Oxford Cancer Centre daily linac checks.

no further than verifying that all the tests had passed, only needing to look deeper to investigate the root cause of any failures. After the testing summary, the selected items of metadata are listed and a copy of the image if requested.

Each of the analysis groups in the tree is then considered in turn, with the outputs of all the analysis modules in a group being collected together. If no tests were performed on any of the results within a group, then the results are presented in a simple tabular format, with each column containing a single type of output. At the bottom of each column, the mean and standard deviation of the results in that column are displayed. The formatting of these aggregate results is controlled by settings in the Report Generator dialogue box. If tests have been designated for any results in an analysis group, then a different tabular format is applied and no aggregate results are calculated. Instead, the name of each analysis module and output is listed alongside columns containing the expected result of the test, the actual result which was calculated, and whether the result was a 'PASS' or 'FAIL'.

At the end of the report is a summary line indicating when the analysis was

performed and the version of IQWorks used, information important for overall quality assurance if the report is to be stored for future reference.

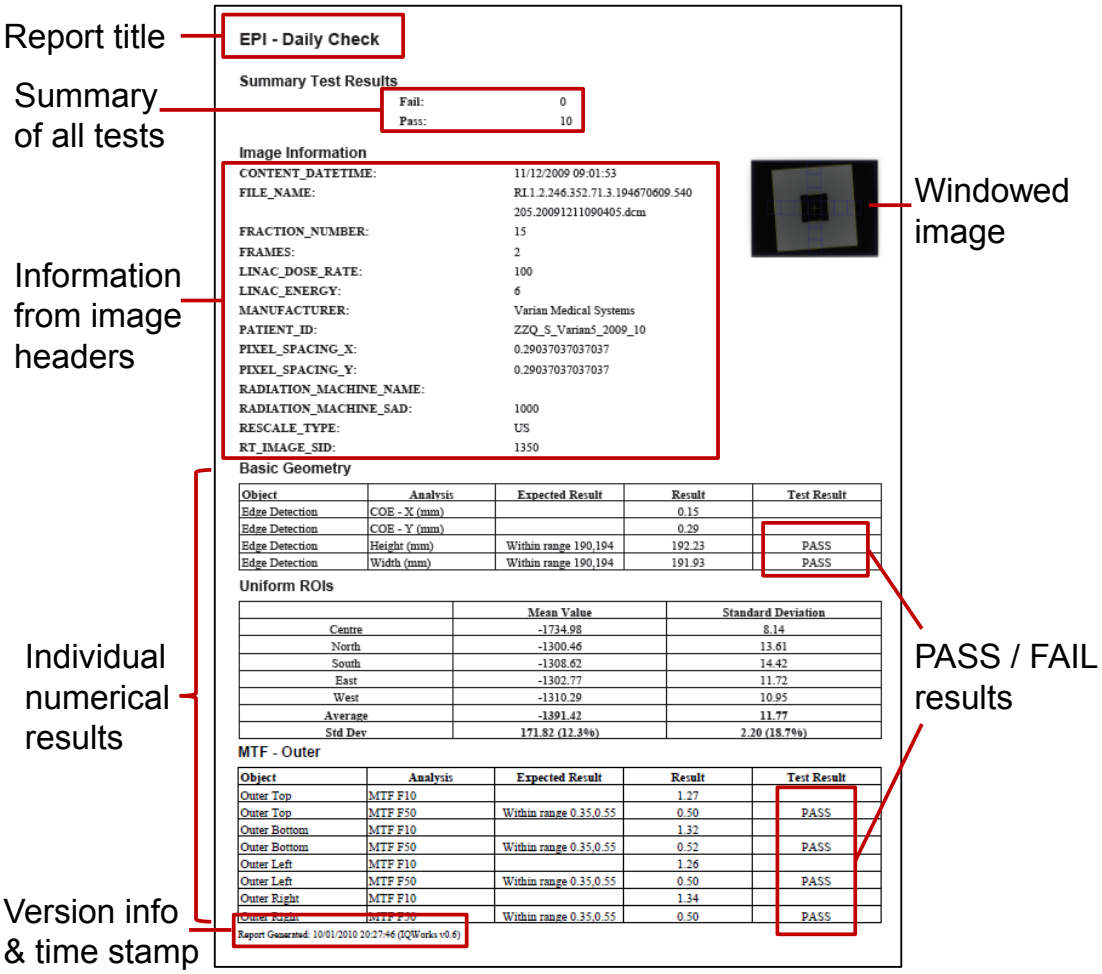


Figure A.8. – Extract of the IQWorks report generated by the Report Generator in the previous figure.

As described above, an analysis tree can be run automatically against multiple images, one after the other. In this situation, apart from an overall summary of all test results being presented at the top, the combined report is constructed by generating individual reports for each image and appending these together.

A.10. Storage of Results

Long term storage of image analysis results is important for many applications, such as the investigation of time trends, forming the historical record of a QA programme and keeping track of a series of experiments. Although IQWorks reports can be printed out or filed electronically more powerful methods are

provided for storing data. All numerical and textual information presented in a report, including associated test results, can also either be appended to a Comma Separated Variable (CSV) file or stored in a database table, with the information to store and the storage destination being chosen via the same Report Generator dialogue box. In the IQWorks analysis workflow, reports can be considered a portal for quickly checking the results of running an analysis tree before the data are stored permanently. Alongside the basic contents of the report, IQWorks also stores a time stamp of when the analysis was performed, the identity of the user who undertook the analysis and an optional user comment. The user name and comment are entered via a dialogue box which automatically appears when the system is ready to store data.

CSV files are convenient vehicles for storing data because they can readily be opened by spreadsheet applications, such as Microsoft Excel. However, once more than a few columns of information are included, or when many rows of results are present, they become unwieldy and difficult to navigate. Instead, relational databases are a more flexible means of storing and querying large amounts of data. To simplify the storage process, both the CSV writing code and the database storage routines inherit from the same generalised storage model. Therefore, regardless of the nature of the storage destination itself the IQWorks reporting code performs similar function calls to prepare the destination to receive information, verify it is ready to receive, send the dataset and to verify it has been received successfully.

Before writing to CSV files, IQWorks automatically checks that the file exists and creates a new one if necessary, including a first row containing descriptive column headers. Data elements are then stored under the appropriate column heading, with empty elements being added if items are missing. Each subsequent addition to a CSV file writes a complete new line of information.

Database storage is achieved via a model which encapsulates simple database interfacing functionality and which must be inherited from and extended to target any specific database platform. IQWorks includes an interface for Microsoft SQL Server which can store data in all editions of this system, both free and commercial, from version 2005 onwards. However, adapting this to target other database platforms based on the Standard Query Language (SQL), such as PostgreSQL or MySQL, would be very straightforward. SQL Server was chosen because it was felt this was most likely to have widespread acceptance in a predominantly Windows and particularly Microsoft orientated healthcare environment. Furthermore, as well as being a robust, efficient platform for

storing data, the installation, configuration and ongoing maintenance of SQL Server tends to require lower overheads than many of the alternatives. It is therefore a good system to rely upon for use in a routine QA programme, where time is limited and specialist database management skills may not be readily available.

Before data can be stored in a database the database server must first be added as an IQWorks *data store*. This then becomes accessible to all analysis trees and is preserved between instances of the program. Because all data stores inherit from the same base classes they can be configured via a common graphical user interface. Each time a storage operation is to be performed IQWorks verifies the table exists and that it is of the correct structure, creating a new table or modifying an existing one if necessary. In particular, if the user changes which information elements are to be stored once a data store has already been created – for example, by selecting an additional item of metadata or deciding to no longer store a particular test result – IQWorks will create additional table fields or automatically modify the storage commands accordingly. This flexibility allows databases to adapt to evolving user requirements without requiring reprocessing of images whenever the analysis scheme is fine-tuned.

When numerical results are presented in a report they are deliberately reformatted to make them easier to interpret visually. This can include rounding to a particular number of decimal places and employing scientific notation, depending upon what has been specified by the user. However, the data stored in CSV files or databases are exactly as generated by IQWorks with no additional processing or rounding. This ensures any further analysis is based upon the raw results and thus avoids any compounding of errors due to the data reduction.

Exporting information to long term storage is only sensible if there is a flexible and efficient means of retrieving it again. However, it was deliberately decided against including this functionality in IQWorks because such data mining (or 'business intelligence') is itself an active field and many advanced and well-established tools for this are already freely available. For example, the free Microsoft SQL Server Reporting Services and the open-source JasperReports are both relatively straightforward to use yet are capable of performing complex queries across multiple database tables and presenting the results in high-quality, customisable reports. In addition, both are also able to prepare detailed graphs based on interactively supplied search criteria. IQWorks was intended to be a robust and flexible image analysis package so it was felt the development effort should focus on its core performance areas rather than branch out to cover

functionality already comprehensively provided elsewhere. As an analysis platform, IQWorks provides the means to load and automatically process large sets of images, generate a preliminary report and store all the results in a database. Other tools can then be used to query the database and generate more complex reports, perform detailed trend analyses, etc.

IQWorks is central to the radiotherapy imaging QA programme in Oxford Cancer Centre, with it being used regularly to analyse images of a range of phantoms acquired using different modalities. For most routine measurements, and particularly for quick quality control checks, the usual scenario is that a pre-defined analysis tree will be run against the images, a report generated and a visual inspection of the report performed to confirm all test results have passed and that there are no anomalous results. All the data are then stored to a QA database. For more detailed analysis, interactive reports are developed using Microsoft SQL Server Reporting Services and deployed via a web interface so that users across the department can run them anywhere without installing additional software. This is a powerful and efficient means of making objective imaging performance data readily available for a variety of purposes. In particular it enables responsible staff to immediately assess the current state of any given imaging modality.

For example, figure A.8 contains a report generated to illustrate trends in EPI contrast to noise ratio measured using the QEPI1 performance phantom as part of the daily check. Presented within the Firefox web browser, the user can examine the results over a specified date range and filtered by linac name (as defined by the name of the test patient), beam energy, dose-rate and the number of frames used to acquire the image. Without requiring additional programming, the contents of the report can be rapidly updated whenever additional measurements are performed. In the example presented a step change in performance due to software reinstallation is clearly visible. This is discussed in more detail in chapter 3.

There are sometimes situations where it is desirable to be able to interactively work with IQWorks data at a more fundamental level, without first needing to construct a report and export information to an external datastore. For example, it may be useful to extract the points of an MTF curve for analysis elsewhere. All graphical results generated within IQWorks can be individually copied in their entirety onto the operating system clipboard for immediate pasting into another package. Furthermore, to facilitate examination of internal results which might otherwise be inaccessible, standard routines have been included

in the IQWorks code base for dumping different types of data to disk files.

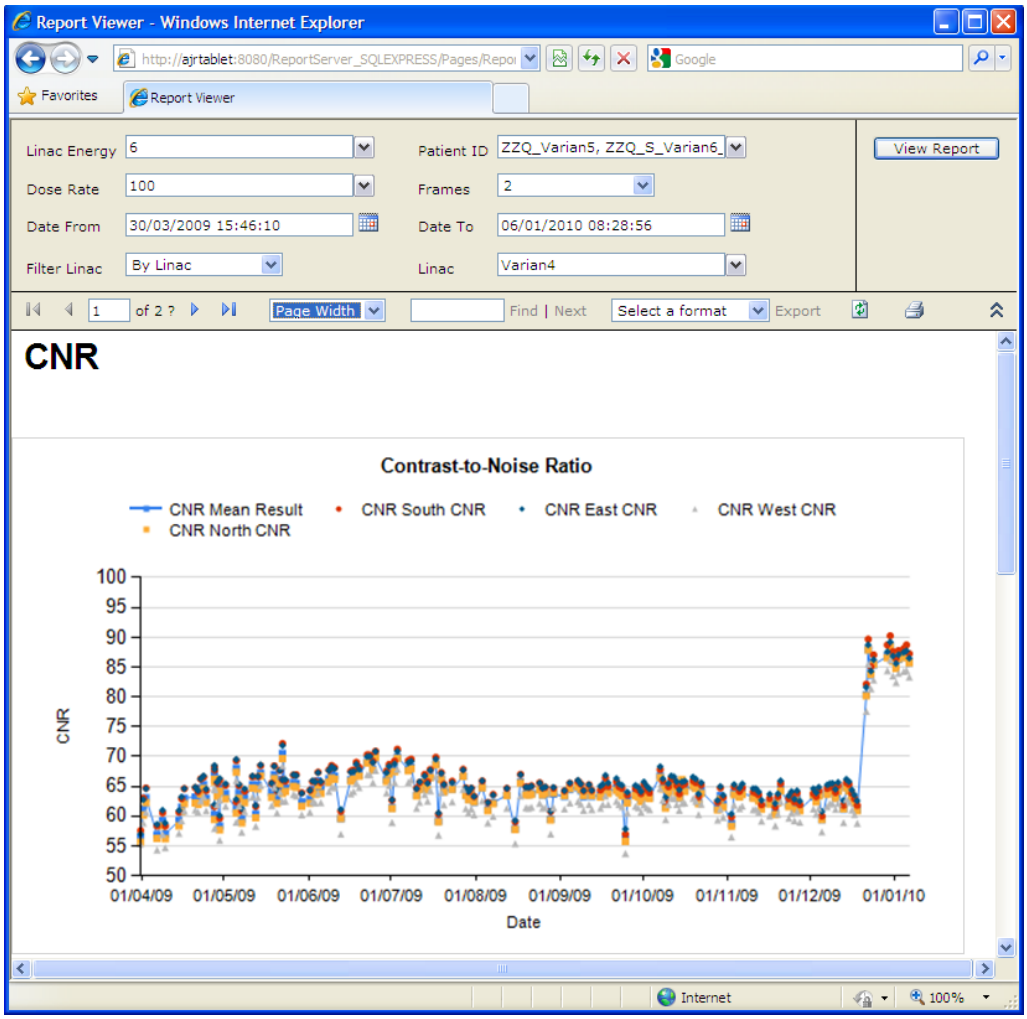


Figure A.9.
CNR time-trend data, calculated and stored to database by IQWorks and presented in a web-browser via Microsoft SQL Server Reporting Services.

A.11. Other Software Frameworks

IQWorks was developed because of a perceived need for an extensible image quality assessment tool which could be used for both research and development, and as part of a comprehensive QA programme. Alternative packages also exist – both free / open-source and commercial – and it is worth comparing these against the IQWorks system. The packages can be divided into two categories: general-purpose analysis frameworks and tools intended for specific applications.

A.11.1. General-Purpose Analysis Frameworks

A.11.1.1. Aquilab ARTISCAN(TM)

Of all the alternative image analysis packages, Aquilab ARTISCAN[11] provides functionality most similar to that of IQWorks. It is a commercial product intended for both radiotherapy and diagnostic applications and includes models of a wide range of phantoms for a number of imaging modalities. It also includes functionality for the reporting and storage of results.

However, ARTISCAN does not follow the 'universal phantom' model introduced in IQWorks, so that although phantoms are analysed via analysis trees these are provided essentially complete by the manufacturer and are not customisable by the user at a fundamental level as in IQWorks. In addition, there is less flexibility in adjusting an existing tree for a new or modified phantom and a new analysis tree cannot be constructed from scratch. Furthermore, because the code is proprietary and closed source, it is not possible to identify exactly how an algorithm has been implemented, or to augment or fine-tune the code base to implement new functionality. This is a major strength of IQWorks.

The developers of ARTISCAN have chosen to integrate reporting, storage and data review functionality within a single package, rather than to separate them logically as in IQWorks. Whilst providing a common package for performing all evaluation and development work, this makes it more difficult to undertake advanced analyses using external software tools or to provide enterprise-wide reporting and review functionality, as described in section A.10.

A.11.1.2. MeVisLab

MeVisLab[367] is a sophisticated visual development environment based around the National Institutes of Health (NIH) Insight Toolkit (ITK)[360]. It therefore has access to a wide range of quality-assured and validated numerical algorithms. Like IQWorks, an analysis scheme is constructed from elemental analysis modules, each of which takes inputs and produces outputs. However, MeVisLab analysis schemes are constructed following a complex graph theory approach which allows more complicated and circular links between modules. Whilst this enables the development of very powerful applications it results in the package being extremely difficult to use, with the interface being complex and inconsistent between modules. Furthermore, MeVisLab has no integrated facilities for generating reports, comparing results against expected values or storing all results to a database. It is therefore more suited to research applica-

tions, rather than as part of a routine QA programme. In addition, full flexibility is only available with the installation of a number of third party dependencies, and through the purchase of a commercial licence.

A strength of MeVisLab is its ability to utilise independently developed plug-ins. However, the 'wrapping' of these in preparation for use is not straightforward and there are inconsistencies in the level of functionality implemented between different modules. A plug-in framework for IQWorks is a high priority for the future, and there is the potential for incorporating and implementing routines from the same NIH libraries utilised by MeVisLab.

A.11.1.3. NIH ImageJ

ImageJ[1] has a strong pedigree, being a Java based successor to the popular Scion Image package. Its development and distribution are managed by the National Institutes of Health, thus assuring the ongoing quality of the software. It is freely available and supports the import of DICOM images by default so is widely used for image processing by the medical physics community.

Although not designed specifically for image quality evaluation, ImageJ implements the necessary Fourier and statistical methods and provides a powerful scripting and plug-in framework through which more complex applications can be developed. Numerous third party plug-ins supporting a wide range of applications are freely downloadable, although the quality of these cannot always be guaranteed. Written in Java, ImageJ runs well on Windows, Linux, Unix and other platforms.

The biggest disadvantage of ImageJ is its limited user interface, which makes development of user-friendly applications for routine clinical use difficult. Furthermore, there is no in-built support for the reporting or storage of results. However, because the ImageJ analysis routines are open-source it is possible to incorporate these into other applications[294], or develop alternative user interfaces and overcome these limitations.

When embarking upon the IQWorks project an extension to ImageJ was seriously considered. However, at the time the UK NHS was experiencing considerable difficulties due to incompatibilities between Java versions and implementations of the underlying Java Runtime Environment. These problems appear to persist today. Also, a number of benefits inherent to the Microsoft .NET framework, including its deeper integration with Microsoft Windows and potential for web-based applications, resulted in the decision not to pursue ImageJ. This said, Java routines can easily be implemented in C# if required,

and Java class libraries even used directly via the IKVM[366] framework, so there is scope for using ImageJ algorithms or even plug-ins in the future.

A.11.1.4. ImPACT+ Calculation Framework

Part of the UK Department of Health's Centre for Evidence based Purchasing (CEP), the ImPACT group has for many years been recognised as world-leaders in the performance evaluation of CT technologies. Members of the team have been responsible for developing an image analysis package capable of performing basic statistical analyses, as well as calculating MTF, NPS and other useful image quality metrics for CT scanners. In particular, the package includes hard-coded routines to automatically process images of the Catphan series of phantoms. Originally intended for CT applications, the calculation framework accepts standard DICOM images as input thereby allowing some of the analysis routines to be applied to images from other modalities. Over time the software has been employed in the rigorous type-testing of numerous CT scanners, with the results being widely disseminated and scrutinised. Its reliability, at least in the field of X-ray CT, is therefore assured and indeed it is generally considered the reference system against which all others must be compared. Although the software is no longer maintained by ImPACT, the original developer and others with access to the code continue to make refinements. However, there is no coordinated, centralised effort and although the software is non-commercial, neither the code nor executable have been freely distributed and they are not readily available to the community. Furthermore, at the end of March 2010 CEP was formally disbanded, leaving the future of ImPACT and the services it provides uncertain.

Another consideration is that the ImPACT calculation framework is written in IDL, a scientific programming language commonly used in the image processing field and familiar to many physicists and engineers[371]. This provides the developer with a comprehensive library of robust analysis routines, all referenced back to basic literature, so it is relatively straightforward to add new functionality to a program without writing algorithms from scratch. However, IDL's support for graphical user interfaces is relatively limited and its programming constructs do not lend themselves to the development of complex, mission-critical applications. Although the compiled software can be run via a freely downloadable IDL runtime environment, modification and recompilation requires an expensive IDL development licence, thus limiting accessibility to the code. Furthermore, the necessity to obtain and install the runtime makes

it more difficult to deploy IDL applications in a clinical environment. There is also no inherent support for document generation to facilitate the reporting of results. Only basic support for data storage is provided by the ImPACT calculation framework as it currently stands. Although IDL provides database interfacing functionality, significant modifications to the program would be required to implement this and it would not be as flexible or customisable as in IQWorks.

It is generally accepted that the ImPACT framework is a valuable research tool but the aspiration is that IQWorks will take its place as the reference tool for CT evaluation. Versions of the ImPACT software maintained by Dr David Sutton (Ninewells Hospital, Dundee, UK) and Mr David Platten (Northampton General Hospital, Northampton, UK) were used in the validation of IQWorks for the analysis of CT and radiographic images.

A.11.1.5. DIMOND3 QA-Distri

Funded by the European Union Sixth Framework programme, the DIMOND3 / SENTINEL project[93] had as its goal the investigation of the link between measurable image quality metrics and clinical performance, aiming to facilitate compliance with European Euratom directives[83]. Part of this remit involved developing methods for calculating objective indices of image quality and a number of software tools were developed, mainly closed-source proprietary packages for use by the project collaborators. However, the 'QA-Distri' set of tools was released by one of the collaborators as an ImageJ plug-in. These are capable of calculating MTF, NPS and DQE and although the algorithms are relatively basic the code is open-source so it is possible to scrutinise them and develop them further if required. Despite this, with the DIMOND3 project now complete the 'QA-Distri' package is no longer actively maintained. Furthermore, because it relies on ImageJ it also suffers from the same limitations.

However, because of the considerable experience of the contributors to the DIMOND3 project, the resulting software can be taken as a baseline standard for radiographic imaging assessment, much like the ImPACT software for CT. It is therefore used in the validation of IQWorks for modalities and phantoms supported by the software.

A.11.1.6. OBJ_IQ_reduced

Acquiring X-ray mammograms as part of a breast screening programme is a delicate balance between keeping doses as low as reasonably practicable whilst maintaining image quality sufficient for the clinical application, particularly so because of the inherent risk in inducing cancer in otherwise healthy patients. As a result of this a great deal of physics input has always been invested in the quality assurance of breast imaging. With the increasing availability of direct digital mammography systems it was recognised that, along with the standardisation of testing protocols, it was necessary to provide reference analysis software to process the digital images. In April 2009, the 'OBJ_IQ_reduced' package was released by the NHS Breast Screening Programme for this purpose[217]. This flexible package accepts DICOM images and provides a wide range of image analysis routines, including MTF, NPS, DQE, variance maps, SNR, CNR and other basic statistical analyses. Although intended for use in mammography it includes modelling of the same system characteristic curves as IQWorks so is applicable to other radiographic modalities.

Written in IDL, the limitations described above in section A.11.1.4 also apply to OBJ_IQ_reduced so that it better considered a research tool appropriate for a dedicated analysis workstation or laptop rather than a package suitable for integration into a live clinical environment. Furthermore, the source code is not freely available so it is difficult to examine exactly how the algorithms work or to implement new functionality. Finally, the primary developer has recently taken up a new position so future support and development of the package is uncertain.

OBJ_IQ_reduced is currently considered the 'gold standard' image analysis tool for mammography so care has been taken to ensure results for IQWorks agree well with those generated by this package. In addition, efforts have been made to include all core functionality implemented by OBJ_IQ_reduced into IQWorks.

A.11.2. Tools for Specific Applications

A.11.2.1. The IRIS Inc AutoQA Lite

'AutoQA Lite'[369] is a commercial product sold by the Institute for Radiological Sciences (IRIS) Inc. It accepts CT images being transferred as DICOM files across the network and automatically analyses images of the Catphan series of phantoms. Intended to be an integral component of an imaging quality

assurance programme it includes reporting functionality, automatically stores results in a local database and allows the user to review results to consider overall performance and aid in identifying trends.

When used exactly as prescribed AutoQA Lite performs reliably and is extremely efficient, analysing a whole series of images in a fully automatic fashion[293]. For this reason it has been licensed by some CT vendors and bundled with their scanners for use as part of their standard QA protocols. However, AutoQA Lite does not perform well when the parameters of the experiment are outside what it is expecting, with it being completely unable to perform an analysis under these conditions. For example, if a particular Catphan slice does not have the longitudinal coordinate expected by the package, or if the phantom is not located where expected in the field of view, or indeed if the field of view is very large so that the proportion occupied by the Catphan section is small, then Auto QA Lite fails. Furthermore, unless the network transfer facility is employed, AutoQA Lite cannot interpret all valid encodings of DICOM files. It is therefore suitable only for basic quality assurance tests in a managed test environment and does not provide scope for more detailed investigations or phantom development.

Although Auto QA Lite incorporates database storage and retrieval functionality this is only to a database stored locally on the computer on which the software is running. It therefore does not scale well as part of a large testing programme where measurements and analyses may be required in multiple locations simultaneously. In addition, the database is not readily accessible to third party tools, making more complex analysis of historical data difficult.

A.11.2.2. Standard Imaging PIPSpro

First developed when electronic portal imaging was an emerging technology, the Portal Imaging Processing System (PIPSpro)[372, 381] has become well established as an image analysis framework providing a range of specialised EPI-related functionality, including matching of EPIs against simulator images or DRRs to evaluate patient set-up errors, analysis of gantry spoke films and EPID dosimetry. PIPSpro also incorporates a powerful scripting language for automating frequently performed activities, thus allowing the package to be utilised as an efficient tool for routine yet complex tasks without forcing the user to adopt a particular approach.

Relevant to this work, PIPSpro has long included routines to analyse the QC3 megavoltage portal imaging phantom[123], with these recently being

extended to cover the new QCkV-1 phantom for the assessment of kilovoltage radiographic imaging. Central to the analysis scheme is the application of Coltman's method for evaluating the MTF from a series of bar patterns[54, 80, 124]. However, unlike the original algorithm PIPSpro normalises the MTF to unity at the spatial frequency of the lowest frequency bar pattern in the test object. This is unfortunate because it yields results which are not directly comparable with those obtained using other MTF calculation methods either in portal imaging or for other modalities. Worse, because PIPSpro presents the results in terms of the reduced metrics f_{50} and f_{10} great care must be taken to avoid these being misinterpreted when they are presented alongside the similarly named results of, for example, an edge analysis. Although the QCkV-1 phantom, along with the PIPSpro analysis scheme, has been warmly received by the radiotherapy community because it has been recognised that objective performance evaluation of kilovoltage radiographic systems is necessary, this phantom / software combination follows a methodology not widely embraced by the diagnostic imaging community. As a commercial package PIPSpro is closed-source and it is not possible to extend its analysis schemes to other phantoms.

For the analysis of CBCT images PIPSpro includes a new module which is an embedded version of AutoQA Lite. This is optimised for the Catphan 504 phantom that is commonly bundled with radiotherapy CBCT systems. However, it is subject to the same limitations of the regular AutoQA Lite package which can be particularly problematic given the relatively poor quality of current CBCT images.

A.11.2.3. PTW epidSoft

This package is intended to process images of the PTW EPID QC phantom[269, 370], although the analysis is only semi-automatic in that the user must first manually identify fiducial markers within the phantom so that the software can place analysis regions appropriately. The software also includes very limited reporting and data storage / review functionality.

Much like the AutoQA Lite package for Catphan CT images, epidSoft is appropriate only for use under well-defined conditions as part of a managed QA strategy. However, its analysis of the PTW EPID QC phantom was compared against that of IQWorks as part of the IQWorks validation process.

A.11.2.4. Quick MTF

'Quick MTF' is a simple package for analysing edge images[373]. Targeting the photographic industry as a lens QA tool it only accepts standard bitmap file formats and does not include functionality beyond generating the MTF from an edge. However, the MTF calculation routines are extremely efficient and robust so that this tool can be relied upon to deliver its specialist functionality when other, possibly more sophisticated packages fail. Although commercial and closed-source, it is relatively inexpensive so was used during the validation of the IQWorks edge analysis routines.

A.11.2.5. Artinis CDRAD Analyser and CDMAM Analyser

For completeness, it is important to include reference here to the Artinis Analyser software[368] for interpreting images of the same manufacturer's CDRAD (radiographic) and CDMAM (mammography) contrast-detail phantoms. These packages implement observer models to determine whether or not details across the surfaces of a phantom are discernible in the image, constructing a contrast-detail curve which can be related back to an ROC analysis. Although it is a long term goal to incorporate this functionality into IQWorks through performing statistical comparisons between the contents of different ROIs this has not been implemented at present because contrast-detail analysis is not fundamental to this current work.

A.12. Conclusion

This appendix presented a new image assessment framework which is applicable across all imaging modalities. It is underpinned by the concept of modelling the 'universal phantom': breaking any phantom for any modality down into its discrete evaluation components. IQWorks was introduced as a software package which implements this framework through automated analysis trees which are constructed from a library of fundamental analysis modules, each of which is designed to target a particular phantom or experiment.

Intended to become a constituent part of a routine imaging quality assurance programme, IQWorks includes automation, reporting, test assessment and data storage functionality. Furthermore, because the development of IQWorks analysis trees is an interactive process achieved via the graphical user interface, constructing new trees from scratch or adapting existing trees is rapid

and straightforward. This ensures that trees can be developed by imaging scientists who may be specialists in their fields but not necessarily advanced programmers, then run automatically by end-users with only basic computing skills. However, by releasing the program code open-source the technical user is also empowered to augment the system through incorporating new analysis algorithms, with the aspiration being that experts will contribute their work so that the package will improve in quality and remain relevant and applicable into the future. Furthermore, making the source code accessible ensures a physicist responsible for image quality evaluation can verify that any given algorithm is performing exactly as intended.

For the first time, IQWorks provides a platform which allows images from both radiotherapy and diagnostic imaging modalities to be considered together and analysed using a common scientific approach and in a common environment. Given the continual convergence of the two disciplines, and the increasing application in radiotherapy of images not originally intended for this, it is hoped IQWorks will facilitate both a deeper understanding of the objective assessment of image quality and the closer working together of professionals from traditionally separate disciplines. IQWorks is already widely used for key image quality measurements, as evidenced by the example in figure A.8 and those presented in chapters 3–5. Indeed, the QA programme for radiotherapy imaging in the new Oxford Cancer Centre relies heavily upon this package as part of a wider strategy for image optimisation.

Appendix B.

IQWorks 2 — Implementation of Assessment Algorithms

B.1. Overview

In the previous appendix the IQWorks image analysis framework was introduced and the philosophy of modelling the 'universal phantom' via analysis trees was discussed. This appendix describes the individual elemental analysis modules which can be used to construct an analysis tree.

Following an object-orientated approach, all IQWorks analysis modules implement the base *IProcess* interface and hence all have fundamentally the same intrinsic functionality. The *IProcess* interface ensures each module includes certain standard functions to enable the IQWorks user interface code to automatically populate the interactive properties dialogue box, associate modules with an image so that they have access to geometrical and pixel data, and display graphical and other results as appropriate. In the same way, additional standard interfaces are also implemented by those modules which expect interaction from the user via the mouse (for example, when resizing or placing a region of interest) or which draw overlying structures on top of an image (for example, to illustrate the line along which a profile is being extracted). Structural elements within the IQWorks framework are also able to automatically identify which code classes 'are' IProcess modules and can therefore make assumptions about the functionality they expose, without requiring prior knowledge of the details of their internal coding. For example, when analysis trees are stored to disk they are stored as text files in Extensible Markup Language (XML) format where the tree structure and every facet of configuration of the individual modules are described in detail. However, because the storage and retrieval routines are able to identify the common structure of an IProcess object

they can automatically deconstruct and rebuild individual modules without additional storage and retrieval routines being developed specifically for them. Using standard interfaces to link analysis modules into the IQWorks framework allows the module developer to concentrate on the core scientific tasks being performed by the module, without being unduly concerned about its interaction with the rest of the framework.

All modules may consume *inputs* from analysis modules earlier in the tree, accept user specified *configuration parameters*, generate *outputs* and may *effect* a permanent change to the underlying image data. This is illustrated schematically in figure B.1 and all IQWorks analysis modules are specified in terms of these characteristics throughout this appendix.

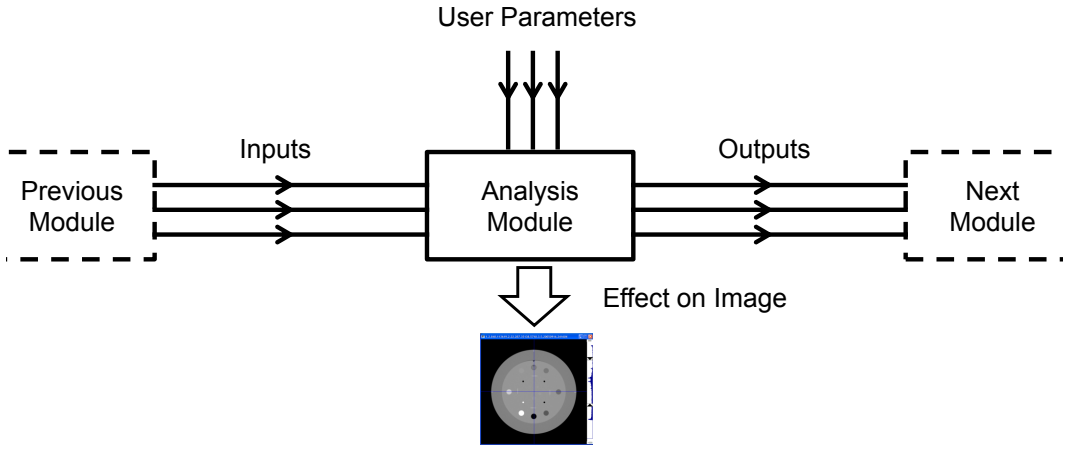


Figure B.1. – Schematic diagram of an analysis module, illustrating its interaction with the user, image and other modules in the analysis tree.

Inputs and outputs can be simple numerical or textual values, more complex structures which encapsulate collections of data, or indeed complete analysis modules themselves. For each anticipated input an object type is specified and only available outputs of the correct type are offered to the user as possible choices when constructing an analysis tree. Configuration parameters are set by the user via the graphical user interface. Often, depending upon the chosen settings of certain parameters other parameters or inputs may or may not be required, or may have different meanings. For example, the 'Arithmetic' module performs simple arithmetic operations on numerical results output by previous analysis modules. If the configuration parameters are set so that two inputs are multiplied together, then the source of both inputs must be provided. Alternatively, if a result is to be multiplied by a user-specified constant, then only one input is required and the constant must be entered as a parameter.

Usually analysis modules work with the underlying image data – both geometrical information and pixel values – but do not actually modify them, for this would involve passing a changed image to subsequent analysis modules. Furthermore, if a tree was run more than once on the same image then the image data would be modified each time so that different results would be generated between runs. However, in certain circumstances it is advantageous to be able to permanently modify the underlying image information, such as in the case of the 'General Fixer' described below. Therefore, all modules intrinsically have access to the underlying data but only a few purposefully change it.

Each atomic analysis module is intended to robustly implement a particular scientific algorithm. However, for many analyses the first steps of data preparation are essentially the same, with only the final stages of the analysis being different. Writing similar functionality repeatedly into different analysis modules is very inefficient and prone to error, resulting in redundant code which is very difficult to maintain. It is far better to develop base classes which encapsulate core processing routines then extend these as required through object-orientated inheritance. In the same way that all analysis modules implement the IProcess interface, many modules inherit from and extend more fundamental modules. When inheriting from another module, the new module assumes all functionality and input, output and parametric characteristics of the base module. Therefore, in the descriptions of individual modules below it can be assumed that a module has all the same characteristics of any module it inherits from, in addition to those specific to itself. Where relevant, hierarchical inheritance trees are presented to illustrate the relationship between different modules.

In section A.5 the importance of code QA and algorithmic validation was discussed. Validation of individual IQWorks analysis modules was achieved through a combination of intercomparisons with other image analysis packages, by referring back to underlying theory and practical experiments. For all but the most simple analysis modules the results of the validation process are presented alongside the module descriptions below.

B.2. General Fixer

Any individual analysis tree is intended to model a particular phantom, with the parameters of its constituent analysis modules being fine-tuned to the specifics of the imaging modality concerned. Ideally, an analysis tree will be applicable

across all instances of a modality without modification, or if additional tuning is required, with only minor changes being made to the configuration parameters. However, sometimes peculiarities of individual instances of a modality yield an otherwise compatible analysis tree inapplicable. For example, some CT scanners do not encode the location of the origin correctly in the image metadata, making it difficult to correctly place regions of interest. Alternatively, pixel size may be unavailable or incorrectly stored, or indeed may have to be adjusted depending upon the experimental set-up. For example, pixel size may be represented in the image metadata in the plane of the detector, rather than at the isocentre plane.

Adjusting the underlying image data to account for modality or experimental anomalies is performed using the 'General Fixer' analysis module, the specifications of which are in table B.1. Unlike most analysis modules, the General Fixer is intended to be run at the very start of an analysis tree and it permanently modifies the underlying image data accessible by other analysis modules.

Inherits from: None			
Inputs	Parameters	Outputs	Effect on Image
None	Adjust origin? Origin – X Origin – Y Adjust pixel size? Pixel size – X Pixel size – Y Auto-window?	None	Adjust pixel size Adjust origin location Auto-window image

Table B.1. – Specification of 'General Fixer' analysis module.

Depending upon the user's choice of parameters, this analysis module can set a new location for the origin, a new pixel size in either matrix direction or can force an image to be auto-windowed to ensure a consistent bitmap display in reports. If adjusting the origin the new location can be specified at either a particular (x, y) pixel location, or to be at the geometrical centre of the FOV. Although similar functionality is provided by the global origin location override described in section A.7, the General Fixer implementation is more flexible in that it only acts when required and can be tailored and reproducibly applied on a tree by tree basis.

B.3. Regions of Interest

Modules to analyse a localised region of interest (ROI) serve two purposes in IQWorks: to enable basic statistical calculations to be performed on the calibrated pixel values within the ROI, and to define the geometrical context within which other image analysis algorithms will act. ROIs are therefore central to almost all useful image processing operations, even if a ROI covers the whole image itself. Four types of ROI are provided in IQWorks: rectangular and elliptical regions, and annular versions of these, with the relationship between these shown in the class hierarchy in figure B.2 and the area covered by each in figure B.3. Essentially, the annular ROIs are the same as their 'solid' counterparts except that a concentric region about the centre is excluded from the analysis. Rectangular and elliptical ROIs tend to be used to extract information from particular test objects within a phantom, such as the mean Hounsfield Units within the PTFE insert of the Catphan, or to calculate the contrast to noise ratio between uniform regions of different attenuation in the QEPI1 portal imaging phantom. On the other hand, annular ROIs may be positioned to surround their solid equivalent, thus yielding information about the region immediately beyond the central portion. They can therefore be utilised for background correction or in statistical decision making where the pixel values within the central and outer ROIs are compared to decide whether or not the test detail would be discernible by a human observer. Whilst subject to the limitations of available observer models, these methods are widely used in the diagnostic imaging community to automatically score images of the CDRAD and CDMAM phantoms[266, 368] and they have also successfully been applied to the analysis of EPI Las Vegas phantoms[77, 324].

All ROIs inherit from the abstract PROI base class which defines functionality common to all ROI modules, including the calculation of statistics, determining whether the ROI lies fully within the field of view of the image and region visualisation tools. Using PROI as the parent class of all ROIs allows other modules to consume ROI statistical information without requiring knowledge of exactly which type of ROI was involved. Internally, each ROI type is represented by a bounding rectangular 'seek region' within which the actual region being considered lies, with the different ROI classes containing code to identify whether or not particular pixels within this larger region should be included in the analysis. A more technical representation of the ROI modules is presented in figure B.4. This illustrates the seek region, the dimension parameters which the user sets (including the X and Y extent of annular components) and the

'handles' at the corners and edges of the seek region for the user to click on and move with the mouse to resize ROIs via the graphical user interface. Modelling all regions of interest within a larger seek rectangle greatly simplifies the programming of the user interface whilst enabling more complex ROI shapes to be implemented in the future with only minimal effort.

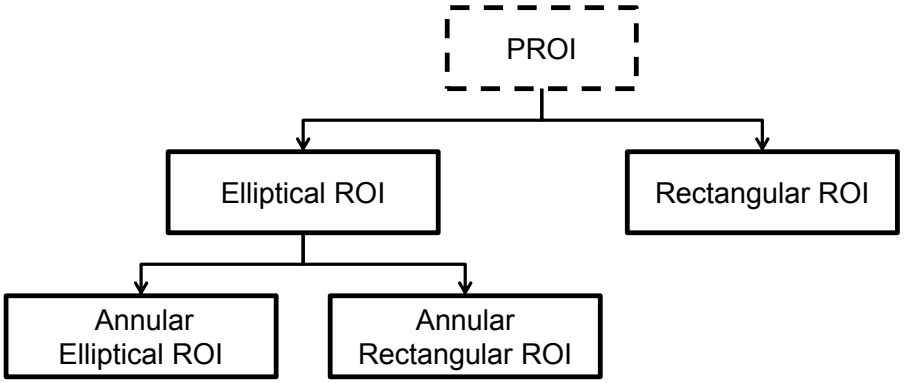


Figure B.2. – Class hierarchy for region of interest types.

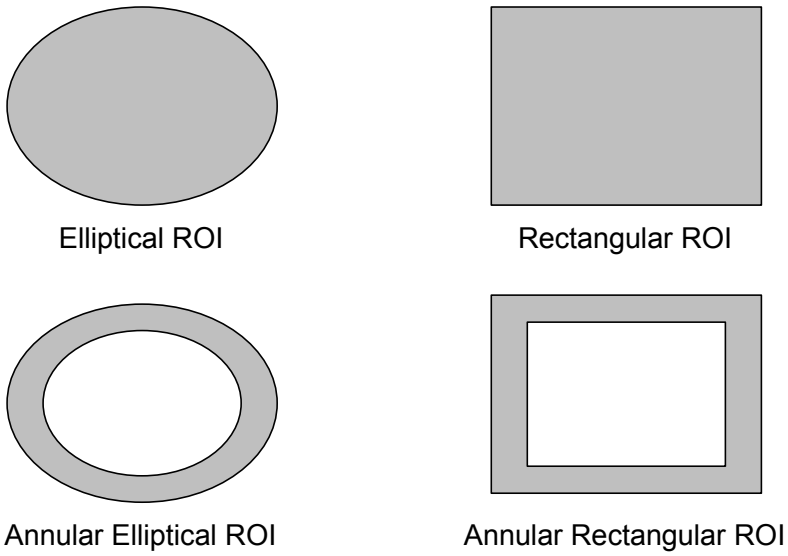


Figure B.3. – Areas considered by different region of interest (ROI) types.

At first inspection it may seem odd that the annular rectangular ROI inherits from the elliptical ROI, rather than the annular rectangular. However, it was decided to adopt this approach for a number of reasons. Firstly, a rectangular ROI which is aligned with the image pixel matrix is a subset of the rows and

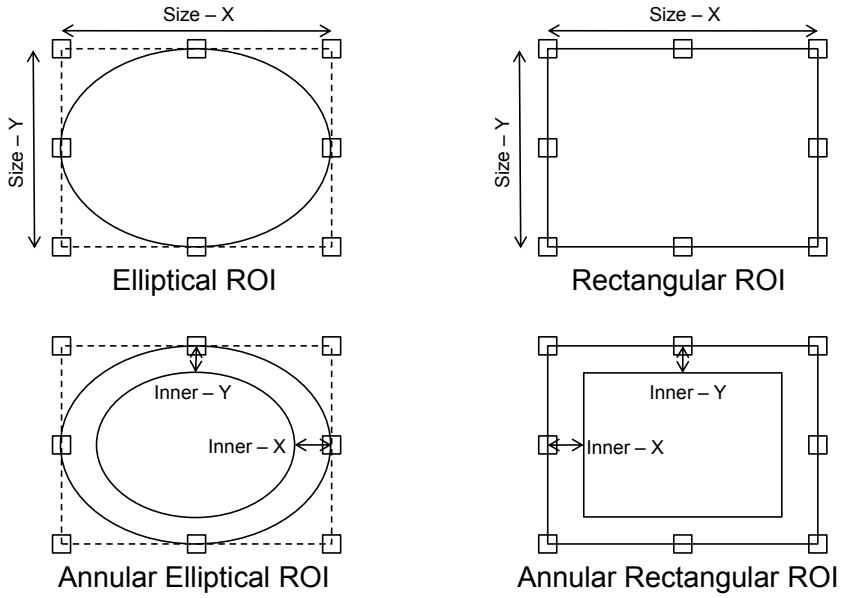


Figure B.4. – Technical region of interest (ROI) specifications. Dimension parameters are indicated and the outer rectangles represent the seek region of each ROI. Interactive resizing of ROIs may be performed by clicking on the square boxes shown at the corners and centres of each rectangle.

columns of the larger matrix so that the pixels within a rectangular ROI can be extracted and analysed by stepping through the matrix pixel by pixel. In contrast, calculations on all the other types of ROI involve stepping through the individual pixels within the seek region and applying a formula to determine whether each is within the region to be considered, including it in the analysis if so and rejecting it otherwise. Whereas the rectangular ROI is the same size as its bounding seek region, so that all the pixels in the seek region are included, all other ROI types involve decision making and the rejection of values. For efficiency, rather than performing this decision making exercise each time a non-rectangular ROI is addressed, instead a list of all pixel values lying within the ROI, along with their coordinates, is constructed the first time it is processed. This list then represents all pixel values to be included in the analysis and is subsequently addressed directly instead of going back to the surrounding seek region, which may be considerably larger. Because the algorithmic logic underpinning the annular rectangular ROI is similar to that of the elliptical and annular elliptical ROIs, it was felt more appropriate that it belonged to this family, rather than inherit directly from the rectangular ROI.

Another reason for not inheriting from the rectangular ROI is that numerous other analysis modules expect to act upon a rectangular region. Some of these

have themselves been implemented by inheriting from the rectangular ROI, where the simple geometry of a rectangular region is assumed. A particular special case is the ROIs used for frequency domain analysis which are required to be simple $[n \times n]$ square matrices, where n is a power of 2 so that the Fast Fourier Transform algorithm can be applied. If more complex ROI shapes were also sub-classes of the rectangular ROI then there would be confusion over which retained the simplistic rectangular geometry and which required special consideration. Therefore, the left hand-side of the tree in figure B.2 is focused purely on statistical calculations within arbitrary ROI shapes, whilst the right-hand side is concerned with both statistical calculations and more advanced analyses.

A final reason for separating the rectangular ROI from the others is that this allows the others to be extended to provide additional functionality, but without breaking the inheritance relationships already in place between other analysis modules and the rectangular ROI. In particular, although all ROIs can be translated and resized they cannot currently be rotated and it is intended to add this functionality to the elliptical ROI and it sub-classes in the future, but not the basic rectangular ROI. Because of the organisation of the class hierarchy this will be possible without creating new classes and without damaging existing analysis trees. Furthermore, no functionality will be lost by not giving the rectangular ROI rotational capabilities because the elliptical rectangular ROI essentially mimics a regular rectangular ROI if its inner X and Y sizes are set to zero.

Inputs	Parameters	Outputs	Effect on Image
Reference Point	Coordinate System Anchor Position Anchor Point Anchor – X Anchor – Y Size – X Size – Y Inner Size – X* Inner Size – Y* Layer Approach Layer Number Pre-filter Surface Plot? Contour Plot? Plot Samples Contour Levels	Statistics Mean Value Min Value Max Value Standard Deviation Variance Processed Point Centre Point Min Point Max Point ROI Object	None

Table B.2. – Specification of region of interest (ROI) modules. * denotes annular ROIs only.

Multi-layer images were discussed in section A.7. When an ROI analysis module is applied to a multi-layered image its behaviour depends upon the value of the 'Layer Approach' parameter and the options for this are described in table B.3. By default the parameter is set to 'Current Layer' so that when a user is working interactively with an image via the graphical user interface the summary statistics results displayed in the properties dialogue box will correspond to the layer currently visible.

Value	Meaning
Current Layer	Analyse the layer currently visible in the graphical user interface.
First Layer	Analyse the first layer in the stack of images.
Specific Layer	Analyse the layer indicated by the 'Layer Number' parameter.
All Layers	Calculate global statistics for all layers in the stack.

Table B.3. – Possible values for the 'Layer Approach' parameter.

IQWorks provides considerable flexibility over the placement of ROI objects so that an analysis tree can be constructed to match user requirements. ROI location and extent can be specified according to a number of coordinate systems, all describing the same frame of reference. All coordinate systems increase from left to right and top to bottom across the image matrix. Different analysis modules within the same analysis tree do not all need to use the same coordinate system.

Taking into account real-world pixel size and the origin of coordinates the 'Physical' coordinate system represents any point in an image as a two dimensional (x, y) vector from the origin, with the magnitude of each component being in millimetres. This system is utilised by default because measurements on medical images tend to be reported in terms of real-world dimensions. Alternatively, the 'Pixels' coordinate system has its origin at the top-left most pixel and dimensions are in units of pixels. However, it is important to note that pixel-based dimensions are not limited to an integral numbers of pixels so that locations can be specified, and the results of measurements reported, with sub-pixel precision. Another system which can be employed is the 'Fractional' coordinate system, in which points are described relative to the physical origin but are specified as a fraction of the field of view. For example, the point $(0.0, 0.5)$ is located at the origin in the x direction and halfway down the field of view in the y direction. Sometimes it is desirable for a region of interest to extend across almost the whole surface of an image, regardless of pixel size, the location of the origin or the dimensions of the image matrix. Usually, the outermost pixels are excluded due to edge artefacts. Both the 'Fractional FOV' and

'FOV Trimmed' coordinate systems yield an ROI centred on the field of view and extending symmetrically by a specified amount either side of the primary axes. Whereas 'Fractional FOV' requires the user to specify the fraction of the field of view to be covered by the ROI in each direction, 'FOV Trimmed' starts with an ROI covering the entire field of view then shrinks it by the specified number of pixels in each direction.

ROI location is specified by 'Anchor – X' and 'Anchor – Y' parameters, and its dimensions by 'Size – X' and 'Size – Y' parameters, with the precise meaning of these depending upon which coordinate system has been chosen by the user. If the coordinate system is changed after the parameters have been set then these are automatically converted to spatially equivalent ones in the new coordinate system, and graphical user interface is updated with the new values. The ROI 'Anchor' describes the coordinates of a point relative either to the origin of coordinates (of whatever coordinate system has been chosen), or to a reference point output by another analysis module. Generally, the first ROI in an analysis tree is used to determine the location of the phantom, an appropriate reference point is output, then subsequent ROIs are placed with their anchors relative to this reference point. Depending upon requirements, the anchor of an ROI can be specified as being at either its geometrical centre or on one of the four corners of its rectangular bounding box. Most frequently, and therefore by default, the Anchor is at the centre of the ROI.

Certain imaging modalities are susceptible to localised point-wise artefacts which may bias statistical calculations or make other analyses difficult. For example, because electronic portal images are relatively low contrast the presence of an uncorrected defective pixel will result in a high contrast spike or dip in signal value that would influence the mean and max / min results calculated in an otherwise uniform region. If the goal was to calculate SNR or contrast then such artificially high or low values would be yield results unrepresentative of the macroscopic situation. To correct for such artefacts it is possible to apply a 'Pre-filter' to the ROI before any calculations are performed. At present IQWorks supports mean or median convolution filters of matrix size $[3 \times 3]$, $[5 \times 5]$ and $[7 \times 7]$. No pre-filter is applied by default.

If the purpose of an ROI is not to provide detailed statistical information about the area, but rather to define a localised region to be acted on by another analysis module, then the statistics calculation is an unnecessary processing burden which can be quite significant if the ROI is large, especially if a pre-filter is involved. An option is therefore provided to suppress detailed statistics

calculations. However, if a child analysis module then requires the statistical information the calculation is turned back on again automatically.

When constructing an analysis tree, and for certain types of analysis, it is often useful to visualise pixel data in formats other than a 2D greyscale image. IQWorks employs R routines to provide facilities for generating contour and surface plots. The 'Plot samples' parameter describes the number of samples of the pixel data which should be chosen in the larger of the two matrix directions, with the number in the other direction being scaled proportionately. If the value of this parameter is large then the generated image will consist of a large number of points and in the case of the surface plot a considerable amount of processing time may be required to draw the wireframe graphic. Once a surface plot has been generated the viewing perspective can be adjusted by the user interactively. For contour plots, the 'Contour Levels' parameter indicates the number of distinct contours which should be included. A surface plot of one of the internal edges of the 'QEPI1' portal imaging phantom is shown in figure B.5, and a contour plot of the whole EPI image in this figure is in figure B.6. This illustrates that the internal and external edges of the lead surface of the phantom are well described by contours at pixel value –1600, suggesting that this value would be suitable for localising the phantom edge in an analysis tree.

B.4. Simple Math

By design, atomic analysis modules yield results of a nature and format appropriate to the image processing algorithms implemented within them. However, practical application of these modules in a real clinical test environment, and particularly their inclusion in a routine QA programme, may require the results of individual modules to be formatted according to a standard protocol or that results from a number of modules be combined into a customised metric utilised in a particular centre. For example, it may be desirable to express the f_{50} calculated by the MTF modules in terms of cycles cm^{-1} rather than the cycles mm^{-1} produced by the modules. Alternatively, the measure of spatial resolution considered most appropriate for a quick daily check may be the average of the f_{50} values calculated from different edges across the field of view, instead of that from any individual edge. Rather than burden atomic analysis modules with these simple arithmetic operations this functionality is provided by the dedicated 'Simple Math' module, the specification of which is given in table B.4.

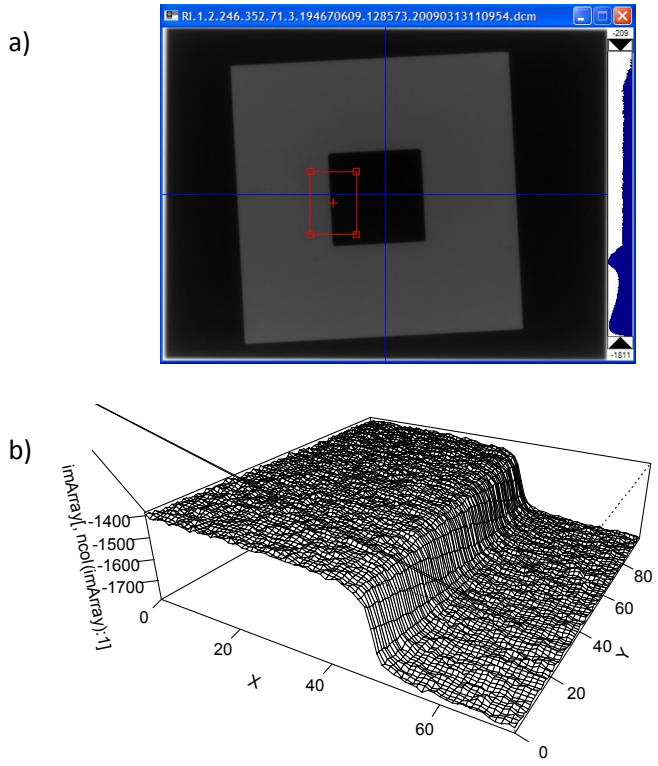


Figure B.5. – Surface plot of one of the internal edges of the 'QEPI1' portal imaging phantom. a) The megavoltage portal image displayed in IQWorks, with the red region indicating the area being considered. b) Surface rendering of the region.

Inherits from: None			
Inputs	Parameters	Outputs	Effect on Image
List of numerical values	Math Operation Use Manual Value? Manual Value	Result Number of Inputs	None

Table B.4. – Specification of 'Simple Math' analysis module.

Simple Math accepts as input a list of any number of numerical values output by other analysis modules earlier in the tree, performs a specified mathematical operation on these, then outputs the result. Mathematical operations available include simple arithmetical calculations (summation or multiplication of the input values, subtraction or division of all values from the first one in the list) and statistical operations (average, standard deviation, variance, maximum, minimum). Depending upon requirements, the user can also supply a constant via the 'Manual Value' parameter. Once the list of numbers has been processed the same operation will be applied to the result of this and the constant. Further operations are also available which expect only a single number to be input via the list. These include taking the square or square root, raising to a specified

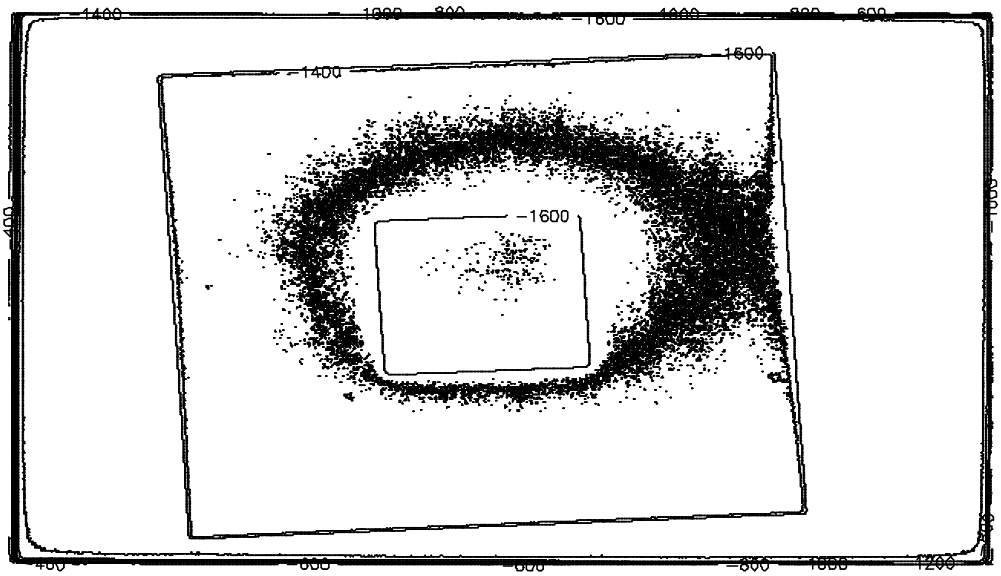


Figure B.6. – Contour plot covering the whole 'QEPI1' image in figure B.5. This plot was constructed using 1024×512 samples and 7 contour levels.

power and finding the absolute magnitude.

In the example applications outlined above, to convert from cycles mm^{-1} to cycles cm^{-1} would involve a simple multiplication by the manual value 10, whilst the average operation would be able to determine the mean f_{50} from an input list.

Results output by Simple Math analysis modules can themselves be input to further Simple Math modules. Complex calculation trees can therefore be developed which encompass the results of any number of analysis modules, effectively allowing the core functionality of IQWorks to be extended by the end user without requiring to resort to programming.

B.5. Edge Detector

Accurate placement of analysis modules requires knowledge of the location of the test objects being considered. Many factors may influence the perceived location of a test object, including the accuracy of laser alignment systems, manufacturing differences between instances of the same model of phantom and differences in geometry between modalities and even instances of the same modality. Unless a phantom can be reproducibly positioned at exactly the

same location in the field of view each time it is imaged it is not possible to reliably place analysis modules based on absolute coordinates. Instead, an 'Edge Detector' analysis module can be employed to find the phantom or test object within the field of view and output reference information which can be utilised to place other analysis regions.

Inheriting directly from the 'Rectangular ROI' module, the 'Edge Detector' analysis module aims to find an edge within the area of the region of interest and characterise this as a list of 'contour' point coordinates. Depending upon the application, the Edge Detector can construct a closed contour which loops back to its starting point (such as from the edge completely surrounding a phantom) or a contour localised to the ROI but from an edge passing through and extending beyond it (such as the edge of a phantom larger than the field of view). A complete specification of the Edge Detector is given in table B.5.

Inherits from: Rectangular ROI			
Inputs	Parameters	Outputs	Effect on Image
None	Detector Type Edge Parameter Closed contour? Deep Search? Contour Map? nth Contour Processed Point Type Area Calc Method	Processed Point COM COE Width Height Contour Map Area	None

Table B.5. – Specification of 'Edge Detector' analysis module. COM and COE refer to Centre of Mass and Centre of Extremes respectively.

When searching for an edge the following algorithm is applied:

1. Extract sub-matrix of image corresponding to the area covered by the rectangular ROI.

2. Apply pre-filter if required. (See section B.3).

3. Run specified edge detection algorithm. This returns a 2D matrix of the same dimensions as in step 1, but with the magnitude at each location being a measure of the 'strength' of the edge there.

4. Apply thresholding to the edge value matrix, resulting in a binary matrix indicating where the edge is present and where it is not.

5. Apply a search algorithm to the binary edge matrix to construct continuous contours from the edge points. When no more points can be found to

continue a contour, a new contour is started. This step in the algorithm returns a list of all contours found within the image.

6. Return the nth longest contour found in the previous step, where n is taken from the 'nth Contour' parameter.

The actual edge detection algorithm applied at step 3 depends upon the setting of the 'Detector Type' parameter. Due to fundamental differences in the operating principles of different modalities it is not possible to obtain a single algorithm which can successfully be employed in all eventualities. However, by giving the user a wide range of options it should be straightforward to adjust an analysis tree to be applicable between different modalities or phantoms simply by modifying the type and parameters of the edge detection algorithm. A description of the edge detection algorithms available in IQWorks is provided in table B.6.

Image segmentation and edge detection are active research fields in themselves. Arriving at a subset of algorithms to provide robust edge detection across all imaging modalities was a considerable challenge. Some of the algorithms implemented, especially those from the CVIPtools library, are sophisticated numerical procedures and a full treatment is outside of the scope of this work. However, what is important is that algorithms are available in IQWorks which allow analysis trees to be constructed where regions of interest may be accurately and reproducibly placed for any imaging modality. The most appropriate edge detection algorithm will vary depending upon the particular modality / phantom combination and in some situations a significant effort may be required by the user to fine-tune its operation before a tree will reliably run automatically. For example, X-ray based modalities tend to yield images with sharp, monotonic edges which lend themselves well to simple thresholding or gradient edge detection. Therefore, the edge of a CT phantom might easily be found by the Threshold algorithm, where the Edge Parameter is chosen as somewhere between -1000 (air) and the mean HU value within the body of the phantom. Alternatively, the edge of a portal imaging phantom might more reliably be located via the Threshold Fraction algorithm, because the range of pixel values may change depending upon the linac beam energy and acquisition settings. This algorithm is also attractive for EPI because it enables standard definitions of the field edge (e.g. 50% of signal level) to be directly applied. However, these algorithms would prove unreliable for segmentation of an MRI image, where sharp edges are susceptible to the Gibb's ringing artefact[238] and the range of pixel values varies considerably between scanners and pulse

Algorithm	Description
Threshold	Find instances where the value specified by the 'Edge Parameter' is crossed over from one pixel to the next. Search by stepping along each row then down each column and combining the results.
Threshold Fraction	As Threshold, but with the 'Edge Parameter' specifying the fraction up the range of pixel values present in the region of interest which should be taken as the threshold parameter. The maximum and minimum values in the ROI are calculated, then the threshold value determined as $\text{threshold} = \min + \text{Edge Param} \times (\max - \min)$
Gradient	Find instances where the difference between the values of adjacent pixels is greater than that specified by the 'Edge Parameter'.
Max Gradient	As Gradient, but automatically choose the maximum difference in adjacent pixel values in any row or column.
Sobel	Apply $\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$ convolution filters separately to the ROI then combine the results of each operation in quadrature. The 'Edge Parameter' is taken as the threshold for converting the resulting edge map from scalar to binary values.
Fuzzy C-Mean Canny	Utilises a routine from the CVIPtools library[338]. Apply a Fuzzy C-Mean segmentation algorithm[206], with the 'Edge Parameter' being the maximum variance permissible before thresholding, then use a Canny edge detector to find continuous edges.
Histogram Threshold Canny	Utilises a routine from the CVIPtools library[338]. Apply a Histogram Thresholding segmentation algorithm[46] then use a Canny edge detector to find continuous edges. The 'Edge Parameter' is not used.
Grey Level Quantisation	Utilises a routine from the CVIPtools library[338]. Divide the range of pixel values into the number of levels specified by the 'Edge Parameter', assign grey levels to the pixels within the different regions, then use a Canny edge detector to find continuous edges.

Table B.6. – Edge detection algorithms available for use in the 'Edge Detector' analysis module.

sequences. Instead, the Fuzzy C-Mean Canny algorithm will generally perform

more reproducibly in such difficult cases.

A detailed comparison of the different edge detection algorithms is not provided here because their performance varies enormously depending upon modality, acquisition settings and nature of the phantom. However, where the Edge Detector analysis module is utilised in chapters 3–5 justification is given of the algorithm chosen each time.

In step 5 of the edge detection process the contour search algorithm applied is:

1. Take a copy of the binary edge map. Each time an edge point is added to a contour it is removed from the copy of the map so that it cannot contribute more than once.

2. Start at point $(0, 0)$ at the top-left corner of the binary edge map.

3. Step along each row until an edge point is found.

4. Examine neighbouring points in the copy of the edge map, looking for another edge point. The search pattern followed is as described in figure B.7. In the figure, square 0 corresponds to the current location in the edge map. By default, only the nearest-neighbour squares 1 and 2 are examined for another edge point, in that order. However, if the edges are not very strong – such as in low contrast or noisy images – then it may be necessary to search over a wider area. If the 'Deep Search' parameter is true then the additional squares 3–5 may also be examined, again in order of increasing distance.

5. When the next edge point is found, remove it from the copy of the edge map and move to that location.

6. Repeat step 4 until no more edge points are found.

7. When no more edge points are found, the current contour is complete. Add it to the running list then go back to step 3 and repeat until there are no more points left to examine.

When the Edge Detector module returns a contour it is drawn in yellow on the surface of the image so that the user can check visually whether it follows a physical edge as expected. The width and height of the contour are calculated and output as the distance, in pixels and physical units, between the X and Y extremes of the contour points. As measures of the geometrical centre of the

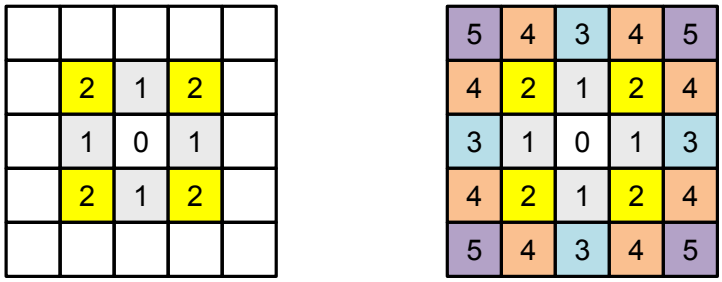


Figure B.7. – Pattern followed when searching for the next edge point to add to a contour. By default, the pattern on the left hand side is used. If 'Deep Search' is enabled, then the additional locations shown on the right hand side are examined.

contour, the centre of mass (CoM) and centre of the extreme values (CoE) are also calculated. Either can be selected to be the location output as the 'Processed Point'. In the event of the edge detection being unreliable, or perhaps when a tree is being run where edge detection is no longer appropriate (such as for a zoomed-in, small field of view image), a failsafe option is included which allows the Processed Point to be defined simply as the centre of the bounding ROI. Whatever is chosen as the Processed Point, this location is marked by large yellow cross-hairs on the image in the user interface. When fine-tuning the edge detection algorithm it can be helpful to visualise the pixels which were identified as edge locations, not just those contributing to the returned contour. If the 'Contour Map?' parameter is activated then an image of the binary edge map is displayed.

An example where the Edge Detector module is employed to localise the QEPI1 portal imaging phantom is shown in figure B.8. In this illustration, a Threshold Fraction algorithm was utilised with an Edge Parameter of 0.5 and Deep Search turned on.

In addition, the Edge Detector module can calculate the area contained within a closed contour. However, this is optional because the operation may be computationally intensive for convoluted contours containing many edge points so must be activate via the 'Area Calc Method' parameter. Two area calculation algorithms have been implemented, both based on work by Cochran[53]. The first involves examining all pixels within the rectangle defined by the contour limits and utilising Franklin's[99] implementation of the Jordan Curve Theorem[117, 129] to count those pixels lying within the contour. Multiplying the number of pixels by the area of each pixel yields the area encompassed by

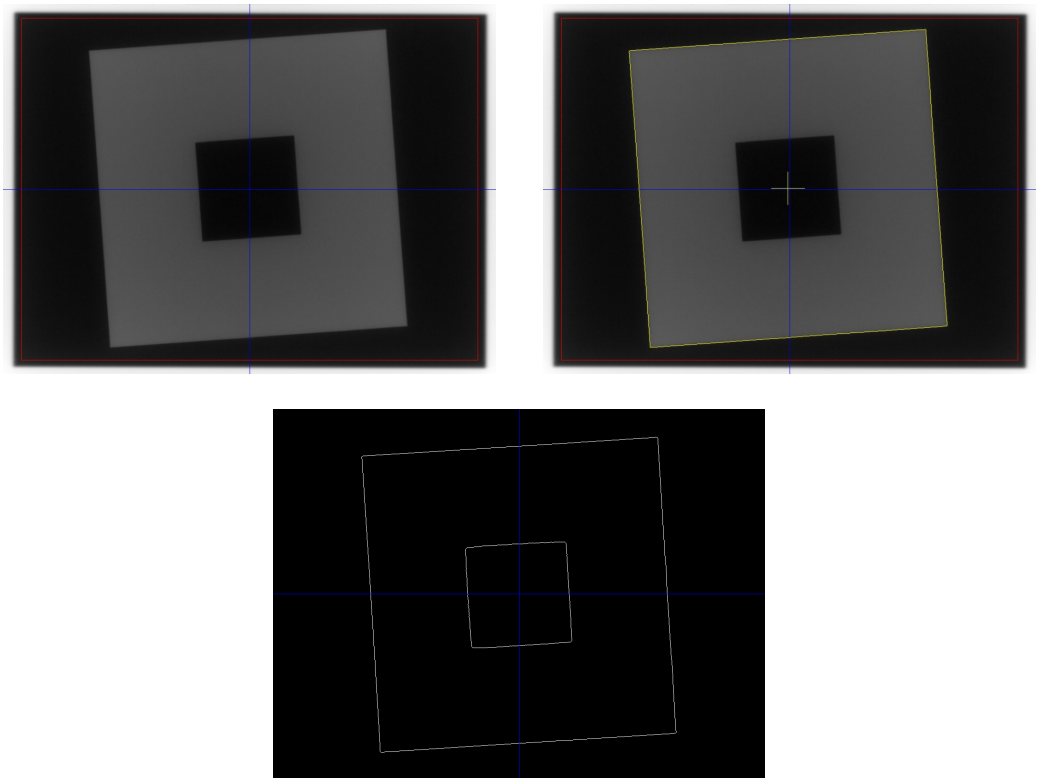


Figure B.8. – Illustration of the Edge Detector module being applied to the QEPI1 portal imaging phantom. Top left: Image before edge detection, with the red outer rectangle indicating the region to search. Top Right: Image following edge detection, where the located edge is indicated by the yellow contour and the centre of extremes by the yellow cross-hairs. Bottom: Binary edge map, showing all edge positions located by the edge detection algorithm. Two contours are clearly visible.

the contour. Alternatively, if the closed contour is considered a polygon with n vertices (x_i, y_i) then the area A can be calculated using the discrete form of Green's theorem[12, 29, 39]:

$$A = \frac{1}{2}\Sigma_{i=0}^{n-1}\left(x_i y_{i+1} - x_{i+1} y_i\right) \tag{B.1}$$

Both methods yield slightly different results, with the pixel-wise approach tending to overestimate the area and the polygonal method underestimating it. From trials the polygonal method is generally faster and produces areas closer in agreement to physical measurements.

Other test object localisation techniques which may be implemented in the future include template matching, based on *a priori* information about the expected shape of the phantom, and more sophisticated segmentation algorithms,

such as the use of geodesic snakes[174]. Although there is scope for improvement, the methods implemented so far have proven reliable and suitable for a wide range of applications, as will be evident in chapters 3–5.

B.6. Distance Measurement

Measurement of distances between reference points in an image is performed by the 'Distance' analysis module, the specification of which is in table B.7.

Inherits from: None			
Inputs	Parameters	Outputs	Effect on Image
Point 1 Point 2	None	Point 1 Point 2 Distance Distance – X Distance – Y Angle	None

Table B.7. – Specification of 'Distance' analysis module.

This module accepts two point objects $P_1\,(x_1, y_1)$ and $P_2\,(x_2, y_2)$ as input and calculates the distance d between them:

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \tag{B.2}$$

In addition, the x and y components and angle θ of the vector $P_1 \rightarrow P_2$ are also calculated

$$d_x = x_2 - x_1 \tag{B.3}$$

$$d_y = y_2 - y_1 \tag{B.4}$$

$$\theta = \tan^{-1}\left(-\frac{d_x}{d_y}\right) \tag{B.5}$$

When the 'Distance' module is executed a blue line is drawn on the surface of the image between the locations of the two reference points so that the user can clearly visualise exactly what is being measured. Furthermore, because the computational burden is relatively light, whenever the reference point locations subsequently change the module automatically reevaluates the inputs and updates the outputs. It is therefore a powerful interactive tool which can be utilised to illustrate the effect of changes in the results of more complex analysis

modules. For example, if the distance is being measured between the CoE points of two 'Edge Detector' modules, then this functionality could be used to evaluate the influence of the edge detection algorithm on the relative locations of the geometric centres of the edges.

B.7. Gradient Analysis

Inheriting from the 'Distance' analysis module, the 'Gradient' module additionally calculates the change in signal value between the two input points as a function of distance. Optionally, the result can be normalised with respect to the signal at a third reference point. This is useful if the gross magnitude of the signal varies significantly between exposure settings or instances of a modality, yet a comparative measure of gradient under standardised conditions is required. The specification of the 'Gradient' module is provided in B.8.

Inherits from: Distance

Inputs	Parameters	Outputs	Effect on Image
Ref Point	Ref Point Norm?	Mean 1 Mean 2 Mean Ref Gradient Gradient – X Gradient – Y	None

Table B.8. – Specification of 'Gradient' analysis module.

Rather than just taking the value of the pixel directly underneath each point, this module discovers the ROI objects R_1 and R_2 used to generate the points, and optionally the normalisation ROI R_{ref}, and calculates the mean signals within each. This method is therefore less influenced by localised image artefacts, such as those due to defective pixels.

Gradients are calculated along each matrix direction, and along the vector between the two points, using

$$G_x = \frac{\overline{R_2} - \overline{R_1}}{d_x} / \overline{R_{ref}} \tag{B.6}$$

$$G_y = \frac{\overline{R_2} - \overline{R_1}}{d_y} / \overline{R_{ref}} \tag{B.7}$$

$$G = \frac{\overline{R_2} - \overline{R_1}}{d} / \overline{R_{ref}} \tag{B.8}$$

where $\overline{R_{Ref}}$ is unity if normalisation is not required.

B.8. Contrast

This analysis module calculates the contrast between two specified regions of interest. Its specification is given in table B.9.

Inherits from: None			
Inputs	Parameters	Outputs	Effect on Image
ROI 1 ROI 2	Calculation Method Make Absolute? Take Log?	ROI 1 – Mean ROI 2 – Mean ROI Difference Overall Mean Contrast	None

Table B.9. – Specification of 'Contrast' analysis module, which acts on two regions of interest (ROIs).

Various definitions of contrast were provided by equations 2.7–2.10 in chapter 2. All of these are implemented by the 'Contrast' module, with the method utilised depending upon the choice of the 'Calculation Method' parameter. In the algorithm the first ROI R_1 is taken as the 'signal' region, and the second R_2 as the background, with differences calculated as $\overline{R_1} - \overline{R_2}$. However, if the sign of the contrast is unimportant then the 'Make Absolute?' parameter can be set to return only the magnitude of the result. Furthermore, for modalities where a logarithmic calculation of contrast is desired, setting the 'Take Log?' parameter will return the base-10 logarithm of the standard result.

The intention of the 'Contrast' module is to provide a comprehensive implementation of all commonly encountered definitions of contrast. This makes it possible to quickly switch between contrast definitions without modifying the wider analysis tree, thus allowing the same intrinsic analysis to be performed across different imaging modalities but with the results still conforming to the generally accepted metrics for each.

B.9. Signal to Noise Ratio (SNR)

Signal to noise ratio (SNR) is defined in equations 2.3 and 2.4. The 'SNR' analysis module is specified in table B.10 and can calculate SNR based on either of these definitions, depending upon the value of the 'Calculation Method' parameter.

Inherits from: None			
Inputs	Parameters	Outputs	Effect on Image
Signal ROI Noise ROI	Calculation Method	Signal Noise SNR	None

Table B.10. – Specification of signal-to-noise ratio ('SNR') analysis module, which acts on two regions of interest (ROIs).

This analysis module takes two ROIs as input – a signal region and a noise region – and it is acceptable for the same ROI to be provided as input for both. In the discussion in chapter 2 it was emphasised that the calculation of SNR should strictly involve only random noise, yet uniform regions on real images typically contain both random and static components. However, it is possible to separate the two by analysing multiple instances of the same imaging scenario acquired under similar acquisition settings.

If there are N layers in a multi-layer image then the deterministic, fixed pattern noise common to all layers is described by R_{det}, which is the average of the ROIs on each layer. i.e.

$$R_{det}(x,y) = \frac{\Sigma_{i=1}^{N} R_i(x,y)}{N} \tag{B.9}$$

where $R_i(x,y)$ is the ROI on layer i. This can then be subtracted from the noise region on any given layer to yield an ROI containing contributions solely from random noise:

$$R_{random,i} = R_i(x,y) - R_{det}(x,y) \tag{B.10}$$

Exactly how this analysis module calculates noise statistics depends both upon the value of the 'Calculation Method' parameter and whether the image supplied is multi-layered.

B.10. Contrast to Noise Ratio (CNR)

Contrast to noise ratio (CNR) is calculated by the 'CNR' analysis module. This is specified in table B.11 and provides numerical implementations of equations 2.3 and 2.4.

The behaviour of this module is similar to the 'SNR' analysis module except that it expects the results of a 'Contrast' analysis module as input rather than a signal ROI. Its treatment of noise, including the removal of fixed-pattern noise from multi-layered images, is identical to that of the 'SNR' module.

Inherits from: None			
Inputs	Parameters	Outputs	Effect on Image
Contrast Object Noise ROI	Calculation Method	Contrast Noise CNR	None

Table B.11. – Specification of the contrast-to-noise ratio ('CNR') analysis module, which acts on a Contrast object and a region of interest (ROI).

B.11. Statistical Comparison

Making statistical comparisons between the pixels in regions of interest to determine whether a test detail can be discriminated was discussed briefly in sections 2.7, A.11.2.5 and B.3. In particular, the potential of using annular ROIs to represent background regions adjacent to their regular counterparts was outlined. As described in section A.4, many of the numerical routines within IQWorks are provided by the *R* numerical analysis package which itself was originally developed to be a sophisticated statistical modelling environment. IQWorks therefore has access to the rich range of statistical comparison methods implemented by R. These are leveraged by the 'Statistical Comparison' analysis module, the specifications of which are in table B.12.

Inherits from: None			
Inputs	Parameters	Outputs	Effect on Image
Signal ROI Background ROI	Statistical Test Confidence Level	p-Value Result	None

Table B.12. – Specification of the 'Statistical Comparison' analysis module, which acts on two regions of interest (ROIs).

This module accepts two regions of interest as input and performs a specified 'Statistical Test' to determine whether the pixels in each belong to the same distribution of values. The level of confidence in the 'null hypothesis' (i.e. the probability that the two sets of pixel values are actually two samples taken from the same distribution) is output as the 'p-Value'. However, when performing a contrast-detail experiment the user is generally more concerned with simply whether or not one can confidently discriminate a test detail, rather than the statistical details. Therefore, to facilitate interpretation of the results the user can also enter a limiting threshold confidence as the 'Confidence Level' parameter. If the p-Value calculated is less than or equal to this then the 'Result' output is TRUE, if not then it is FALSE. In the literature[59, 77] 5% is often selected as the threshold confidence level and so this value has been chosen as the default.

Three statistical tests have been implemented in IQWorks, the technical details of which are outside the scope of this work:

- The Student *t*-test[59], where it is assumed the pixel values within both ROIs are normally distributed, with both distributions having the same variance. Although frequently utilised in the literature this test is limited because the assumption of normal distribution may not be valid, depending upon the nature of the imaging modality and test object. Furthermore, unless the test conditions are very low contrast, so that the signal and background ROIs have very similar means, then even if the normality assumption is true the variances of the two distributions will be unlikely to be the same.

- The Wilcoxon Rank Sum Test[215, 352], which is similar to the Student *t*-test but without the assumption that the samples are normally distributed.

- The Kolmogorov-Smirnov Test[216], which employs cumulative probability distributions to compare various characteristics of the two samples (mean, variance, skew, etc.) and generates a single comparative metric.

For any particular scenario the most appropriate choice of test will depend upon the nature of the imaging modality and design of the experiment. A common theme in the literature is that a great deal of fine-tuning, experimentation and indeed an element of trial and error is required to arrive at a statistical test which matches the performance of a human observer, even under well-defined, controlled experimental conditions[266, 324, 334].

B.12. Profile Analysis

An image 'profile' describes signal value as a function of position along a straight line lying either within a single two-dimensional image plane or across a stack of slices in a volumetric dataset. The 'Profile' analysis module described here handles profile extraction from a two-dimensional plane, whilst the 'Stack-Profile' module, covered in section B.14, deals with stacks of images. A full specification of the 'Profile' module is included in table B.13.

Internally, the 'Profile' module maintains a simple $y = mx + c$ model of a straight line. However, there is considerable flexibility over how this line is defined, with the intention being that a single analysis module should comprehensively be able to accommodate all practical eventualities. The most

Inherits from: None			
Inputs	Parameters	Outputs	Effect on Image
Reference ROI Point 1 Point 2	Profile type Profile length Centre point type Angle Background Correction Profiles to average Invert profile? Force positive? Calculate FWHM? Plot profile?	Profile points Background mean FWHM Centre point Centre of FWHM Point 1 Point 2 Statistics Uniformity – Integral Uniformity – Integral+ Uniformity – Integral- Uniformity – Differential Coefficient of Variation	None

Table B.13. – Specification of the 'Profile' analysis module. FWHM refers to the calculation of the Full Width of the profile function at Half its Maximum intensity.

straightforward method of specifying the line is to provide its start and end points: 'Point 1' and 'Point 2' in the table. These might originate from other analysis modules, or can be placed manually. For example, it may be desirable to extract a profile between the origin of coordinates and an aluminium rod embedded in an otherwise tissue-equivalent CT phantom. In this situation, the origin would be specified manually whilst the centre of the rod would be the coordinates of the maximum pixel value in a ROI surrounding the rod.

An alternative method is to provide the length of the profile line either side of a specified centre point in a reference region of interest, along with the angle of the profile (in degrees) relative to the vertical axis. Similarly to the previous method, the centre point can be either placed manually or an output from the reference ROI (e.g. location of maximum pixel value). In addition to being able to enter a numerical value for the angle directly, there is also a drop-down list of the most frequently chosen directions, including the horizontal and vertical and the positive and negative diagonals. Also, instead of specifying an absolute length it is possible to choose for the profile to extend across the limits of the ROI. If a profile is specified using this method then the start and end points of the line are calculated internally and are output as 'Point 1' and 'Point 2' respectively.

Providing such flexibility has the side-effect that a number of the input parameters are interrelated. When this is the case dependent parameters automatically update in real time when others are modified. For example, if the

profile direction is changed to be 'Horizontal' then the angle automatically displays 90°.

Once the profile characteristics have been specified a blue line is drawn on top of the image so the user can identify the path followed. If parameters are subsequently changed – for example 'Point 1' or 'Point 2' is moved – then the line updates in real-time to reflect the new specification.

Profile values are determined by stepping pixel by pixel across the image matrix in either the X or Y direction, calculating the coordinates of the point where the line intersects the normal to the stepping direction at each location, identifying the pixel lying at that point, then assigning the pixel value to the profile at that distance. The stepping direction is chosen to minimise the step size depending upon the angle of the line. For example, if the profile line is only a few degrees from the vertical then the Y matrix direction is followed. If the profile direction is horizontal or vertical then the returned value is average of a specified number of adjacent rows or columns, otherwise the raw pixel value along the line is returned.

A background region can also be defined as a specified length either side of the limits of the profile. If so, then the mean pixel value is calculated in this region – being the mean of the means either side of the profile – then this is subtracted as a baseline correction from each of the points in the profile.

If required, the profile can be automatically inverted to force a positive absolute maximum value following any baseline correction. Another option is to simply always invert the original profile.

Following calculation of the profile the points can be output to another analysis module, or the results plotted graphically for visual inspection or data extraction. The full-width at half maximum (FWHM) can be calculated if desired, with the full-width value and location of its centre provided as module outputs. If the FWHM option is selected then the geometrical construct is also plotted on top of the curve itself. Other module outputs include the value used for background correction, standard statistics from the profile points (mean, standard deviation, max, min, etc.) and the uniformity measures defined by equations 2.14 to 2.18.

B.13. Slice Thickness

Tomographic slice thickness can be determined from an image through a ramp test object by measuring the FWHM of the profile across the ramp and taking

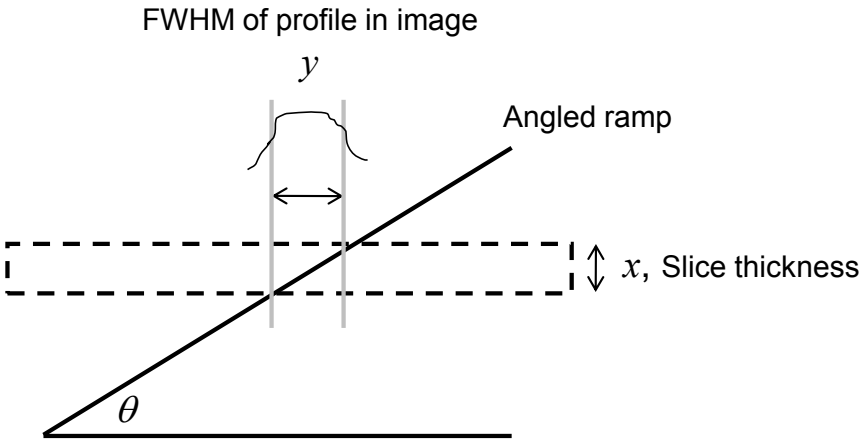


Figure B.9. – Schematic diagram illustrating the calculation of slice thickness from the image of an angled ramp.

into account the ramp's angle. If a ramp at angle θ is intersected by a slice that produces an image in which the FHWM of the apparent profile of the ramp is measured as y, then the slice thickness x is given by

$$x = y \tan \theta \tag{B.11}$$

as is illustrated in figure B.9.

This module is specified in table B.14 and accepts the output of an existing 'Profile' module as input, along with a specified ramp angle, then uses equation B.11 to calculate the effective slice thickness from the two.

Inherits from: None			
Inputs	Parameters	Outputs	Effect on Image
Profile	Edge Angle, θ (degrees)	Slice Thickness	None

Table B.14. – Specification of the 'Slice Thickness' analysis module.

B.14. Stack Profile

This module, specified in table B.15, is functionally equivalent to the 'Profile' module described above. It generates exactly the same numerical and graphical outputs as that module but operates on a stack of images rather than a two-dimensional plane. Its method for calculating the profile is also different. Rather than representing the profile internally by a model of a straight line, each plane in the image is assigned one profile point at a distance coordinate determined

by the image's position in the stack. The value of the profile at each plane is then taken from a nominated point within a reference ROI. This can be specified as any standard output point type, such as the location of the maximum or minimum pixel value within the region, or just simply the coordinates of its centre point.

If requested, background correction is performed by subtracting the mean pixel values within ROIs in a specified number of slices at the start and end of the stack.

Inherits from: None			
Inputs	Parameters	Outputs	Effect on Image
Reference ROI	Profile Type Profile Length Background Correction Region to Average Invert Profile? Force Positive? Calculate FWHM? Plot profile?	Profile Points Background Mean FWHM Centre Point Centre of FWHM Statistics Uniformity – Integral Uniformity – Integral+ Uniformity – Integral- Uniformity – Differential Coefficient of Variation	None

Table B.15. – Specification of the 'Stack Profile' analysis module. FWHM refers to the calculation of the Full Width of the profile function at Half its Maximum intensity.

B.15. Uniformity Analysis

Numerous uniformity metrics are calculated by the 'Profile', 'Stack Profile' and 'ROI' modules described above. However, the dedicated 'Uniformity ROIs' module, specified in table B.16, calculates uniformity by considering gross differences in mean pixel values between two groups of ROIs. If one of the groups contains ROIs at the centre of the field of view, and the other ROIs at the periphery, then the results of this module are compatible with the uniformity metrics defined for CT in IPEM Report 32[156] and by Varian Medical Systems in the performance specifications of their OBI CBCT system[345].

Apart from the last three, the outputs listed in table B.16 are self-explanatory.

In addition, the 'Noise' output is the mean standard deviation across all ROIs,

and is calculated using:

$$Noise = \frac{\sum_{i=0}^{N_{Central}} \sigma_i + \sum_{j=0}^{N_{Peripheral}} \sigma_j}{N_{Central} + N_{Peripheral}} \tag{B.12}$$

where $N_{Central}$ and $N_{Peripheral}$ are the numbers of regions of interest in the central and peripheral ROI groups and σ_a is the standard deviation in ROI a. If divided by 10 then equation B.12 is equivalent to the CT noise metric in IPEM Report 32, where the average standard deviation is essentially divided by the mean value for water (1000 HU in the offset HU scale) and multiplied by 100 to give a result as a percentage. This can easily be implemented directly in an analysis tree using the 'Simple Math' module if required.

U1 and U2 are uniformity metrics calculated by

$$U1 = \frac{m_{Central} - m_{Peripheral}}{m_{Central}} \tag{B.13}$$

$$U2 = \frac{m_{Central} - m_{Peripheral}}{10} \tag{B.14}$$

where $m_{Central}$ and $m_{Peripheral}$ are the means of the mean pixel values in the the central and peripheral ROI groups respectively. Both are equivalent in that they are the difference in means divided by the mean of the central ROIs. However, whereas U1 is universally applicable (as long as the system is linearised so that the values in a uniformly exposed image cannot be zero or less!) U2 is taken from IPEM Report 32 and assumes a non-linearised CT geometry in which $m_{Central}$ on the denominator is replaced by 1000 (the HU of water in the offset HU scale) and the result multiplied 100 to become a percentage.

B.16. Variance Map

This module generates a new image with pixel values representing the result of a specified statistical calculation within a range about each pixel in the original image. Although it is named 'Variance Map' the module can also calculate local standard deviation and local SNR (in terms of local mean divided by local standard deviation) in addition to the local variance.

Variance (or similar) images of uniformly exposed flood fields are useful for highlighting localised detector artefacts[217, 218]. Typically the range is set to only 1 or 2 pixels either side of each reference pixel so that only neighbouring

Inherits from: None			
Inputs	Parameters	Outputs	Effect on Image
ROI Group 1 (Central) ROI Group 2 (Peripheral)	None	Max Difference Between Means Min Difference Between Means Global Mean Global Std Dev Group 1 – Mean of Means Group 1 – Std Dev Group 1 – Minimum Mean Group 1 – Maximum Mean Group 2 – Mean of Means Group 2 – Std Dev Group 2 – Minimum Mean Group 2 – Maximum Mean Noise U1 U2	None

Table B.16. – Specification of the 'Uniformity ROIs' analysis module. The metrics U1 and U2 are defined in the main text.

pixels of the next-nearest neighbours are considered. Although to an extent these images can be interpreted visually the number of mapped values outside a specified tolerance can also be taken as an objective measure of image quality. For example, a defective pixel may be identified as having a local variance more than two standard deviations greater than the average local variance. By regularly counting the number of mapped pixels outside this tolerance any increase in defective pixels can be tracked over time. Depending upon the application, tolerances can be specified as either an absolute value, a fraction of the mean value in the map, or a particular number of standard deviations. In addition, the user can decide whether the tolerance should by applied positively only (i.e. whether the value of a pixel must be *below* the mean + tolerance for it to be within spec) or whether it should be positive and negative (i.e. lying within a band either side of the mean).

Sometimes, the values of certain pixels in a map image may be so large that they overwhelm contributions from others which are large relative to a normal signal but are nevertheless very small in comparison with these outliers. This may occur at the edges of a detector, if a test object is present so that it adds structure to an otherwise uniform field, or in a region of a detector where there is already known damage. These excessively large values make it difficult to interpret map images either visually or through counting 'bad' pixels. A facility

is provided to mitigate this problem by ignoring all mapped pixel values above a specified threshold, which by default is set at 1000. Experimentation may be necessary to arrive at the most appropriate threshold for a particular application or modality.

Inherits from: None			
Inputs	Parameters	Outputs	Effect on Image
Reference ROI	Map Type Local Range Truncate Large Values? Truncation Threshold Generate Tolerance Image? Tolerance Value Tolerance Type	Map Mean Map Max Map Min Map Variance Map Std Dev Pixels Outside Tolerance Absolute Tolerance	None, but the variance map and optionally the tolerance map are generated as new images.

Table B.17. – Specification of the 'Variance Map' analysis module, which acts on a rectangular region of interest (ROI).

B.17. Modulation Transfer Function (MTF)

B.17.1. Overview

IQWorks is capable of calculating the MTF from three different types of fundamental test object: from a series of bar patterns of different spatial frequencies; from an angled edge or line object; and from an impulse or point object. Each test object is analysed via a dedicated analysis module and although the algorithms underpinning each module are different the outputs are largely the same. These include the MTF curve, the line or point spread function deduced from the object (if applicable), the applied data windowing function and requested reduced resolution metrics. If required, any output curve data can be plotted within the package for immediate visual analysis.

Available reduced resolution metrics include the spatial frequency f_x at which the MTF falls to a fraction x% of the lowest frequency value in the curve (usually $MTF\,(0)$) – most commonly f_{50} or f_{10}, the 50% and 10% frequencies respectively – and the Noise Equivalent Aperture (NEA). NEA is thought to be a single figure of merit which agrees well with visual perception of sharpness[355] and is calculated using

$$NEA = \left(2\pi \int_0^{f_c} MTF^2 f \mathrm{d}f \right)^{-1} \tag{B.15}$$

where f_c is the Nyquist Frequency.

B.17.2. MTF from a Series of Bar Patterns

As detailed in chapter 2, the Modulation Transfer Function describes how the magnitude of a sinusoidal signal of frequency f changes as it passes through a system. In principle, given an appropriate test object containing patterns each containing a single spatial frequency the MTF could be measured directly. Unfortunately, it is extremely difficult to manufacture such test objects and it would is not possible to construct a universal phantom applicable to all imaging conditions. However, square-wave bar patterns consisting of alternating opaque and transparent strips are relatively easy to build. Seminal work by Coltman[54], and refined by Droege and Morin[80], describes the relationship between a sinusoidal output and a bar pattern (square-wave) input. In general, the MTF at a given spatial frequency is proportional to the modulation under a bar pattern of that frequency. Therefore, the MTF curve can be determined through measuring the relative modulation in a series of ROIs placed on bar patterns of different spatial frequency. As long as a sufficiently large number of points are measured over a range appropriate to the device being evaluated then a reliable MTF curve can be constructed. The IQWorks 'Bar Pattern MTF' performs this assessment for a list of ROIs placed on bar patterns of known spatial frequencies and is specified in table B.18.

Inherits from: None			
Inputs	Parameters	Outputs	Effect on Image
List of ROIs Uniform ROI 1 Uniform ROI 2	Spatial Frequency of Each ROI MTF Calculation Method Normalisation Method Aliasing Correction Noise Correction Frequencies f_x	MTF Curve Noise Level List of f_x NEA	None

Table B.18. – Specification of 'Bar Pattern Modulation Transfer Function (MTF)' analysis module, which accepts a number of regions of interest (ROIs) as input. NEA refers to the Noise Equivalent Aperture.

In the underlying theory 'modulation' is defined as the difference in signal value between the bars and gaps, divided by their sum. Usually, this is implemented in code by taking the standard deviation in the pixel values within each ROI, and this is the method IQWorks utilises by default. However, some

software packages attempt to apply the definition directly[275]. An option has therefore been included to calculate modulation as

$$MTF(f) = \frac{max_f - min_f}{max_f + min_f} \tag{B.16}$$

where max_f and min_f are the maximum and minimum pixel values within the ROI of pattern frequency f, although the experience during this work was that this method is relatively unstable due to it being oversensitive to fluctuations in pixel values not caused by the bar patterns themselves (such as the presence of defective pixels or a high background noise level).

Bar patterns are essentially a series of box functions and contain an infinite range of spatial frequencies, albeit present at a lower level than the fundamental frequency due to the periodicity of the bars and gaps. This means that the modulation measured in a ROI on any particular pattern contains contributions from the other frequencies in the pattern. The original Coltman formulation includes a correction for these contributions which involves adding or subtracting from any given ROI the weighted modulations from higher frequency patterns

$$MTF'(f) = MTF(f) + \frac{MTF(3f)}{3} - \frac{MTF(5f)}{5} + \frac{MTF(7f)}{7} - \ldots \tag{B.17}$$

which can be generalised as

$$MTF'(f) = MTF(f) + \Sigma_{n=1,2,3\ldots}(-1)^{n-1}\frac{MTF((2n+1)f)}{2n+1} \tag{B.18}$$

However, this is not universally applied[281] and indeed is sometimes found in a modified form[275]. IQWorks therefore offers the user the choice of no aliasing correction, the pure Coltman correction or any known alternatives, with the default being the pure Coltman.

By convention the MTF should be normalised so that it is unity at zero spatial-frequency. This is important if MTF curves calculated via different algorithms or test objects are to be meaningfully intercompared. Because bar patterns by definition always represent a spatial-frequency greater than zero special treatment is required to calculate a value for the zero-frequency modulation. Droege and Morin show[79, 80] that the relative modulation at zero-frequency can be calculated as half the difference between the mean pixel values in ROIs placed on uniform patches of the same materials which constitute the bars and

gaps in the bar pattern, as long as the regions of interest are sufficiently large that their size has no influence on the mean signal. i.e.

$$MTF\,(0) = \frac{\left|m_{\text{bar pattern}} - m_{\text{gaps}}\right|}{2} \tag{B.19}$$

where $m_{\text{bar pattern}}$ and m_{gaps} are the mean pixel values in the ROIs on the bar and gap materials respectively. However, some researchers do not calculate $MTF\,(0)$ but instead normalise with respect to the relative modulation of the lowest frequency bar pattern in the series[275, 281]. Whilst this is a reasonable approach if performance is always to be measured using a single test object it prevents comparisons against metrics calculated using test objects or algorithms where the normalisation is different. IQWorks defaults to performing a zero-frequency normalisation based on the signal values within two supplied uniform ROIs. However, normalisation against the modulation at the frequency of any of the bar pattern ROIs can be performed instead if required.

Stochastic image noise adds a baseline modulation to all regions an the image which can significantly influence the calculated MTF if the contribution is large relative to that of the bar patterns. A number of noise compensation strategies have been implemented and the user can select to employ any of these or none. All strategies involve calculating a noise variance N^2 and subtracting this from the originally measured modulation:

$$M_{corrected}\,(f) = \sqrt{M^2_{measured} - N^2} \tag{B.20}$$

N^2 can be chosen to be either the average variance in the two specified uniform ROIs; the average variance across all bar pattern ROIs; and either of these in a difference image from which fixed pattern structure and noise should be missing. It should be noted that these strategies are not all equivalent and do yield different results, although the particular method selected by different researchers appears to be somewhat arbitrary and there is little consistency in the radiotherapy imaging literature. However, it is author's experience that if care is taken to minimise the image noise level when undertaking an MTF assessment using bar patterns then the influence of noise is negligible and no correction is required.

If a multi-layered image is supplied as input to this analysis module then the user can choose to generate an MTF curve either from any particular layer or which is the average curve across all layers.

In summary, the algorithm followed by the 'Bar Pattern MTF' analysis module

is as follows:

1. If required, calculate noise variance N^2 using selected method.
2. Calculate standard deviation in the ROI specified for each bar pattern.
3. If required, perform noise correction.
4. If required, perform aliasing correction.
5. Perform normalisation either at $f = 0$ using uniform ROIs and equation B.19, or against the modulation in a specified bar pattern ROI.
6. Calculate request reduced resolution metrics (e.g. f_{50}, NEA, etc.)
7. Return MTF curve

B.17.3. MTF from Edge or Line

Calculation of the presampled MTF from an edge or line test object is performed by the 'Edge Line MTF' analysis module, specified in table B.19. This module aims through parametric configuration to be flexible enough to produce results consistent with the majority of similar algorithms described as being actively used by established research groups in the literature (see, for example Samei and Flynn[307], Marshall[218] and the algorithms compared in the papers by Samei *et al.*[306, 308] and Neitzel *et al.*[258]). In particular, it is fully compatible with the IEC 62220-1 standard for DQE evaluation[146].

Similar steps are followed whether the test object is an edge or a line / slit, except that the line spread function (LSF) is calculated directly if the object is a line, rather than by differentiating an edge spread function. Fundamental to this module is the ability to obtain an oversampled line or edge spread function by analysing an image of a line or edge lying at a shallow angle to one of the pixel matrix directions, usually around 3°. The process is illustrated in figure B.10 for an edge at angle θ to the vertical matrix direction. If the column of numbered pixels is considered then the distance of any pixel from the edge is given by

$$d = x \sin \theta \tag{B.21}$$

An incremental step from pixel to pixel up the column therefore results in a much smaller corresponding change in distance to the edge, allowing a finely sampled ESF to be constructed.

The algorithm implemented by this module is:

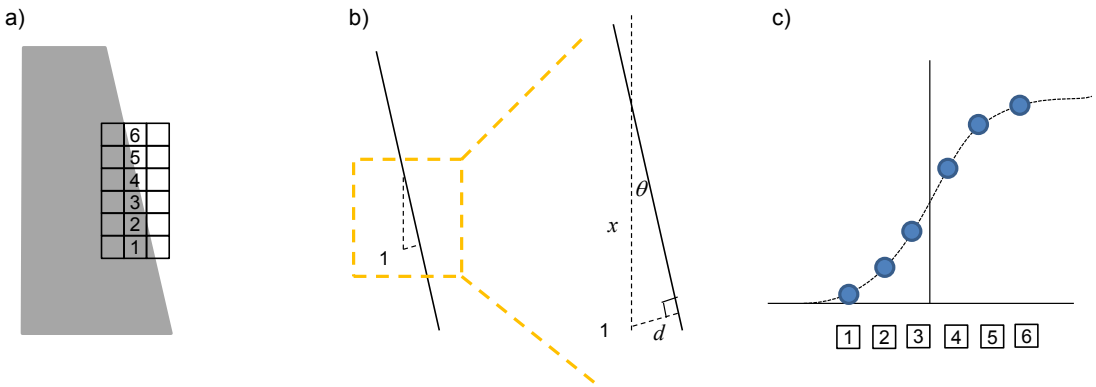


Figure B.10. – Illustration of analysing an angled edge object to obtain an oversampled edge spread function. a) The edge lying on the pixel matrix. b) Geometrical construct for oversampling. c) The oversampled ESF.

Inherits from: Rect ROI

Inputs	Parameters	Outputs	Effect on Image
None	Test Object Type Edge Detection Method Edge Detection Range Sample Size Sample Smoothing No of Points in LSF Part of LSF to Use Data Window MTF Frequency Binning Frequencies f_x	ESF Curve LSF Curve LSF FWHM MTF Curve Edge Equation Edge Angle List of f_x NEA	None

Table B.19. – Specification of 'Edge Line Modulation Transfer Function (MTF)' analysis module. LSF refers to the Line Spread Function, NEA is the Noise Equivalent Aperture and FWHM indicates the Full Width at Half Maximum of the MTF curve.

1. **Determine the equation of the edge or line.** One of a variety of edge/line detection methods can be selected depending upon requirements including a maximum gradient filter, Sobel edge detector, Canny edge detectors or a double Hough transform followed by a Sobel or Canny edge detector. If the seek ROI is large then some edge detectors may be biased by noise or structure far from the edge or line. The user can therefore opt to seek only within a specified range, in which case IQWorks performs a first pass using the whole ROI to roughly localise the edge / line then repeats the process in the narrower band either side of the results of the first pass. The line is modelled internally by a simple $y = mx + c$ equation.

2. **Calculate the Line or Edge Spread Function.** Tabulate pixel value as a function of distance either side of the edge or line, rejecting pixels beyond

a maximum distance if required. The tabulated values are then rebinned to a uniform spacing, with the central bin centred on the edge or line. Bin size can be specified either as an absolute value in millimetre or as a fraction of the pixel size (the 'Sample Factor').

3. **Perform smoothing if required.** If required, a noisy line or edge can be conditioned using a locally fit Gaussian weighted polynomial smoothing algorithm[307, 329].

4. **If the object is an edge, differentiate to yield the Line Spread Function.**

5. **Rebin or interpolate the LSF to a specified number of points.** This must be a power of 2 to facilitate utilising a Fast Fourier Transform algorithm.

6. **Apply a windowing function if required.** Any of the windowing functions in figure 2.6 can be applied over the line or edge sampling range.

7. **Modify the LSF if required.** Usually, an MTF is calculated by taking a Fourier Transform of the whole LSF. However, in some situations – particularly when a very noisy edge has been differentiated – one half of the LSF may be better defined than the other and this may prevent the FFT algorithm from producing a satisfactory result. If it is known that the LSF should be symmetrical then the user can choose to consider only half the LSF, with this being reflected back over the origin to give a symmetrical LSF for input to the FFT. Either the left or right hand side of the LSF can be selected specifically, or the software can try to calculate which of the two yields the better result. By default, IQWorks always uses the whole LSF.

8. **Take the Fourier Transform of the LSF to give the MTF.** IQWorks employs the Fast Fourier Transform implemented by R[379].

9. **Normalise the MTF to unity at zero frequency.**

10. **Rebin or Interpolate the MTF to a specified spacing, if required.**

Because the 'Edge Line MTF' module implements a relatively complex algorithm which it is intended should be compatible with those of many other workers it is important that it be carefully validated. This was achieved by comparing the performance of this module against that of three independent packages when analysing the same reference images: OBJ-IQ_reduced,

DIMOND3 QA-Distri and a customised version of the ImPACT+ software prepared by Mr David Platten, an imaging physicist at Northampton General Hospital, who contributed to earlier releases of this software when previously working for the ImPACT device evaluation group. David Platten has since been involved in the work of the Institute of Physics and Engineering in Medicine's diagnostic radiology special interest group in encouraging the adoption of objective and quantitative analysis of QA images in diagnostic imaging. (Details of the ImPACT+ software can be found in section A.11.1.4 of the previous chapter.) Two datasets were considered in this evaluation: a simulated perfect edge image generated by software (figure B.11) and a digital radiograph of a tungsten edge (figure B.12), both supplied by David Platten. The pair of images was analysed by each of the software packages and the MTF curves resulting are plotted underneath the respective images in the two figures.

From the plots it is evident that there is generally good agreement between all four packages, with the differences between them being of the order expected for this type of evaluation[258, 306, 308]. Indeed, for the real tungsten edge image, there is near perfect agreement between all software implementations. However, for the simulated edge image, whereas OBJ-IQ_reduced and David Platten's results agree well, the MTF curve calculated by QA-Distri is slightly higher. It was discovered that IQWorks could be set to mimic either of these scenarios: if the Sample Factor (SF) was set to 0.145 then the results agreed almost perfectly with OBJ-IQ_reduced and David Platten, if set to 0.05 it was in line with those of QA-Distri. This is an important result because it emphasises the potential differences between supposedly objective metrics depending upon configuration parameters, and it is also encouraging because it demonstrates that IQWorks can mimic established software packages for the purposes of comparison.

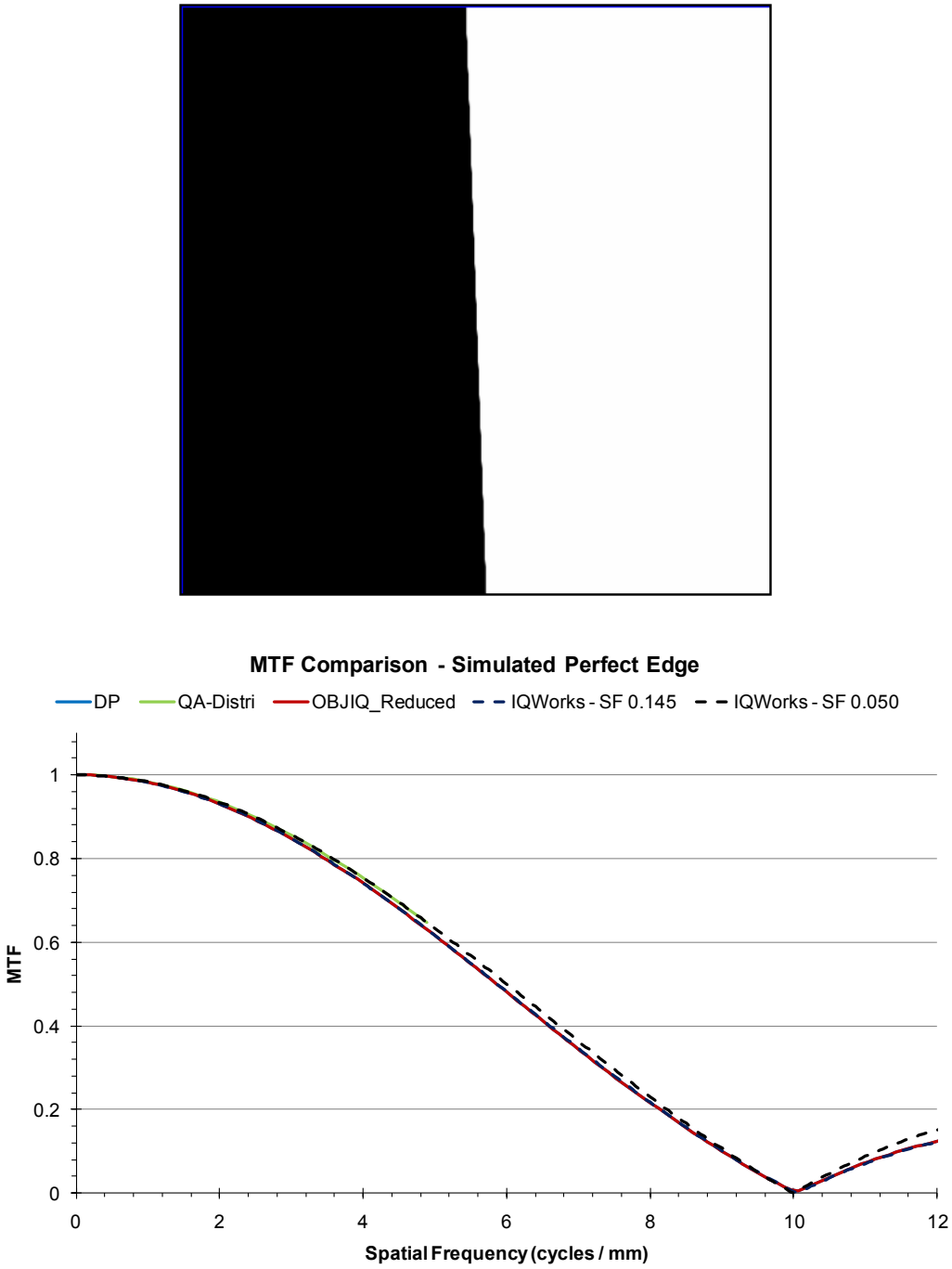


Figure B.11. – Top–Simulated edge image, courtesy of David Platten. Bottom–Modulation Transfer Function (MTF) curves calculated by different analysis packages: David Platten's software (DP), DIMOND3's QA-Distri, OBJ-IQ_Reduced and IQWorks with two Sample Factor (SF) settings.

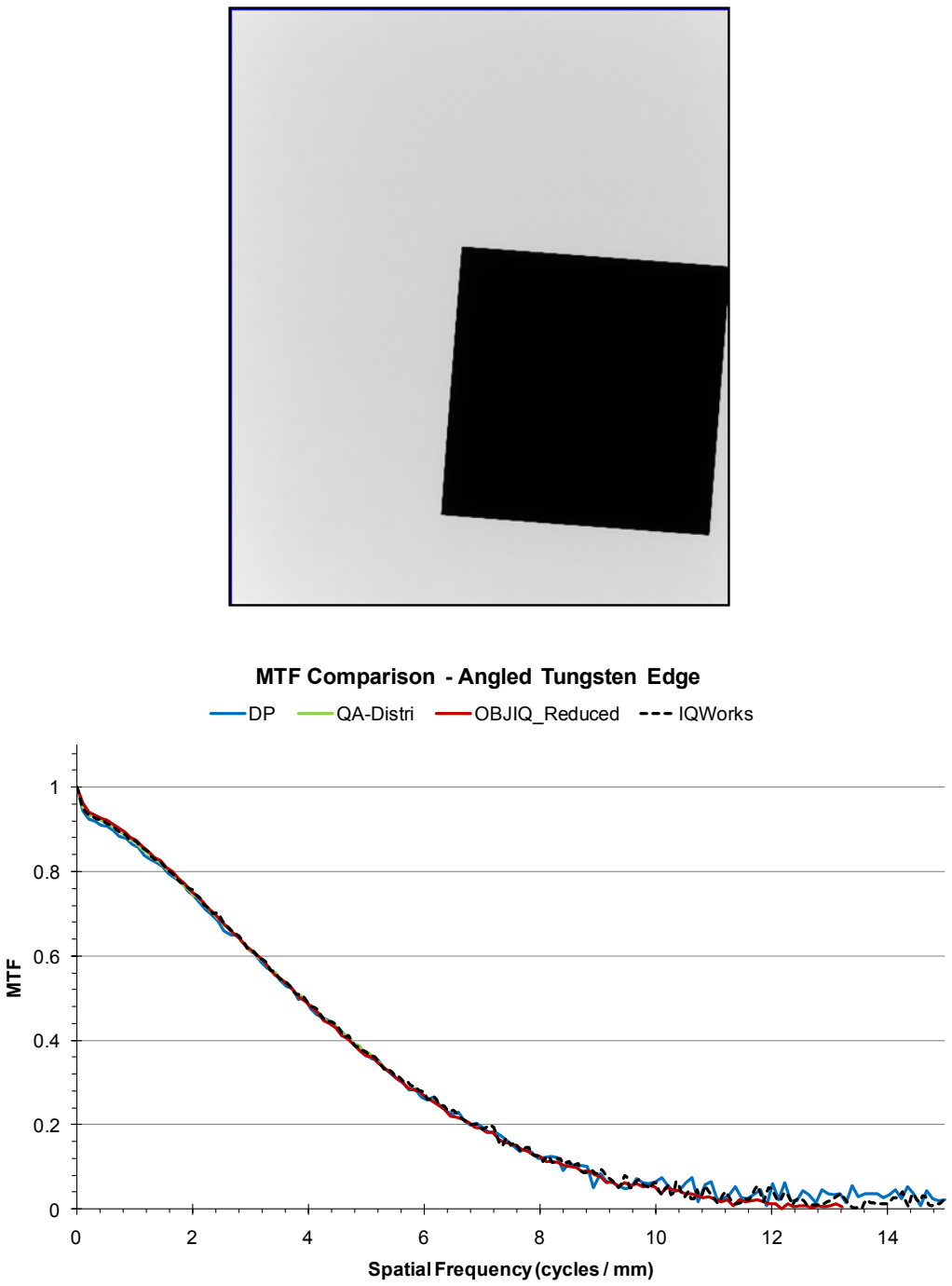


Figure B.12. – Top–Digital radiograph of a tungsten edge, courtesy of David Platten. Bottom–Modulated Transfer Function (MTF) curves calculated by different analysis packages: David Platten's software (DP), DIMOND3 QA-Distri, OBJ-IQ_Reduced and IQWorks.

B.17.4. MTF from Impulse Object

Analysis of an impulse object is performed by the 'Impulse MTF' module, specified in table B.20. This is very similar to the 'Edge Line MTF' module except that the Point Spread Function (PSF) is calculated in the first stages of the algorithm rather than a LSF.

An impulse object is typically a point-like structure of either high or low signal and the 'Centre Point Type' parameter is used to specify whether the centre of the impulse should be taken as the maximum or minimum pixel value in the seek ROI. If it is expected that the PSF should be radially symmetrical then all pixel values within the seek range are tabulated as a function of radial distance from the centre point, otherwise the average of a specified number of row or column profiles is taken. This behaviour is specified by the 'PSF Type' parameter and the resultant PSF is background corrected using the mean value within an annulus around the seek region if the PSF is radially symmetrical, or rectangular ROIs either side of the profile limits if not. The width of the annulus or rectangular ROIs is set by the 'Background Width' parameter. It can be difficult to manufacture test objects containing impulse structures which are sufficiently small that they do not themselves cause a blurring of the perceived PSF and thus a degradation of the measured MTF. If required, the 'Object Diameter' parameter can be set to deconvolve the calculated MTF by modelling a cylinder of the specified diameter. All other parameters perform similarly to those of the 'Edge Line MTF' module.

Inherits from: Rect ROI

Inputs	Parameters	Outputs	Effect on Image
None	Centre Point Type PSF Type Background Width Sample Range Sample Smoothing No of Points in PSF Part of LSF to Use Data Window Object Diameter MTF Frequency Binning Frequencies f_x	PSF Curve PSF FWHM MTF Curve List of f_x NEA	None

Table B.20. – Specification of 'Impulse modulation transfer function (MTF)' analysis module. PSF refers to the Point Spread Function, NEA is the Noise Equivalent Aperture and FWHM indicates the Full Width at Half Maximum of the MTF curve.

In summary, the 'Impulse MTF' module algorithm is

1. **Calculate the Point Spread Function.** If the PSF is expected to be radially symmetrical then all pixels in the seek region are utilised, otherwise row or column profiles are taken.
2. **Perform smoothing if required.**
3. **Perform background correction.** Uses an annulus if the PSF is radially symmetrical or rectangular ROIs either side of the profile limits if not.
4. **Rebin or interpolate the PSF to a specified number of points.** This must be a power of 2 to facilitate utilising a Fast Fourier Transform algorithm.
5. **Apply a windowing function if required.** Any of the windowing functions in figure 2.6 can be applied over the line or edge sampling range.
6. **Modify the PSF if required.** (See description under the 'Edge Line MTF' module.)
7. **Take the Fourier Transform of the PSF to give the MTF.**
8. **Normalise the MTF to unity at zero frequency.**
9. **Deconvolve the MTF if required.** Divide the MTF by the modulus of the transfer function of a cylinder of specified diameter.
10. **Rebin or Interpolate the MTF to a specified spacing, if required.**

B.18. Noise Power Spectrum (NPS)

B.18.1. Overview

Two NPS analysis modules are provided by IQWorks: 'NPS Multi ROI' and 'NPS Auto ROI'. Both modules perform identical analyses except that whereas 'NPS Multi ROI' accepts a list of arbitrarily placed ROIs the 'NPS Auto ROI' module takes a single ROI and automatically places sub-ROIs within this. The approach taken by 'NPS Auto ROI' is employed almost universally in the literature when calculating the NPS of flat panel detectors, but it is difficult to apply when there is not a clear rectilinear geometry (such as in CT or MRI), where the field of view is of relatively low dimensions or where there are known regions of bad pixels).

Similarly to the 'Edge Line MTF' module described above, the intention was to develop a generic NPS algorithm consistent with any of those found in the

contemporary literature, such as the work by Marshall[218] or the methods reviewed by Neitzel *et al.*[258], Dobbins *et al.[282]* or Ranger *et al.*[282]. Again, particular care was to taken ensuring this module was capable of generating results in line with IEC 62220-1[146].

B.18.2. Core NPS Algorithm

Regardless of the NPS module selected for a particular analysis the core algorithm followed is:

1. **Ensure validity of all ROIs.** Verify that all ROIs lie on the image, are squares of the same size and that the square side is a length which is a power of 2 so that the FFT algorithm can be utilised.

2. **Construct list of ROIs to process.** This simplifies the calculation by making the algorithm more general. If the image is multi-layered then ROIs are taken either from a single specified layer or from the same locations on all layers.

3. **Calculate basic statistics in all ROIs.** Depending upon requirements, these may be used as part of accounting for signal trends or for normalisation purposes.

4. **Perform trend removal from all ROIs..** Before taking the FFT any background signal must be removed so that only fluctuations due to noise are present. IQWorks supports sophisticated trend removal which is performed on a region by region basis. The simplest option is to subtract the mean pixel value in a region from all pixels in that region. Alternatively, planar or 2D polynomial surfaces can be fitted and subtracted. These options are illustrated in figure B.13.

5. **If required, perform normalisation to account for long-distance trends.** Even following regional trend removal there may be residual long-distance trends which can bias results. If required, the noise signals from the previous step can be normalised by multiplying by the ratio between a reference signal level and the mean signal in the original ROIs. The reference signal can either be the mean pixel value in a nominated ROI or the global mean across the whole image.

6. **Perform data windowing.** Any of the window functions in figure 2.6b can be utilised, all of which are normalised to have an RMS of unity so that they do not influence noise power calculations.

7. **Take FFT of each ROI to calculate 2D NPS.**

8. **Perform NNPS normalisation if required.** If calculating NNPS, divide each of the FFT images from the previous step by the square of the reference signal from step 5.

9. **Calculate stochastic and fixed pattern NPS images.** Apply equations 2.72 and 2.73.

10. **Calculate 1D NPS in X and Y matrix directions.** Stochastic and fixed pattern 1D noise power spectra are calculated by considering the pixel values in bands of specified width either side of the X and Y axes on the NPS images from the previous step. Contributions can be excluded from either just the axis being considered or from both axes. The spatial frequency of each NPS point is determined by adding the x and y spatial frequency coordinates in quadrature.

11. **Rebin or interpolate each 1D NPS to the specified interval.**

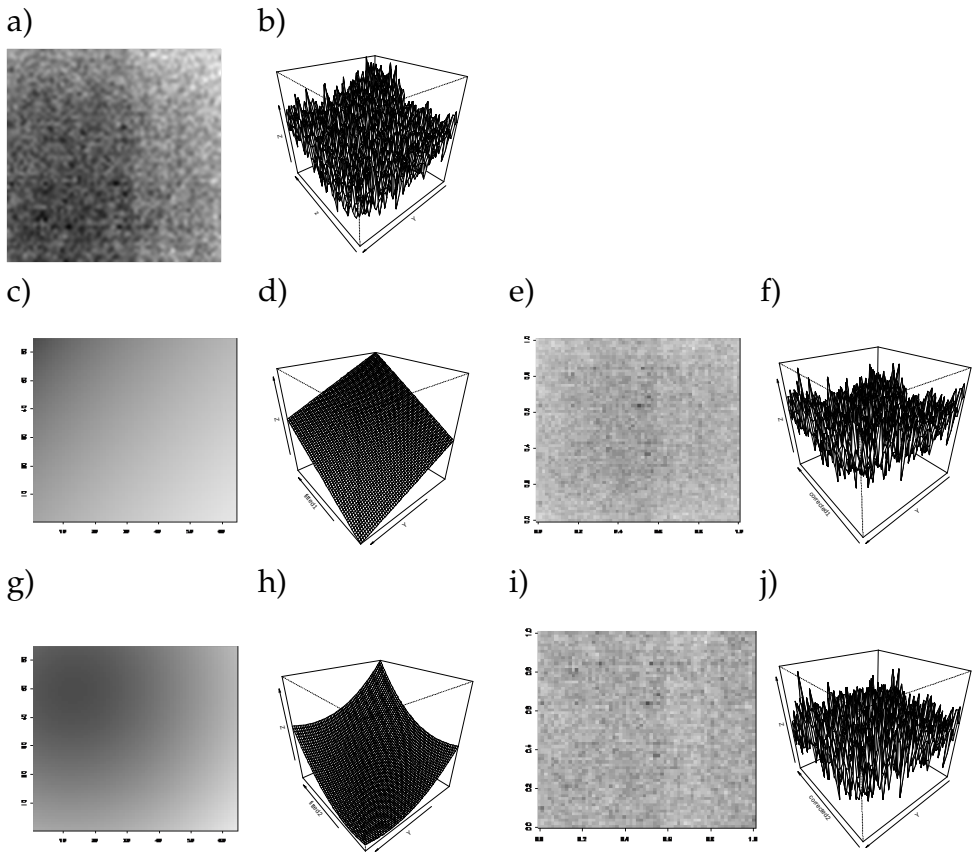


Figure B.13. – Illustration of trend removal algorithms. a) Original region of interest (ROI) and b) its surface plot. c) Fitted plane with d) its surface plot. e) ROI following subtraction by plane in c) and f) its surface plot. g) Fitted 2D second order polynomial to a), with h) its surface plot. i) ROI following subtraction of g) and j) its surface plot.

B.18.3. 'NPS Multi ROI' Module

This is the more general of the NPS calculation modules, being a basic interface to the core NPS algorithm. It is specified in table B.21.

Inherits from: None			
Inputs	Parameters	Outputs	Effect on Image
List of ROIs	Layer Approach Trend Removal Method Normalisation Type Normalisation ROI Data Window Calculate NNPS? Profile Width Axes to Exclude NPS Rebin Interval	Stochastic NPS Image Fixed Pattern NPS Image Intermediate Images 1D NPS Profiles Data Window	None

Table B.21. – Specification of the 'Noise Power Spectrum (NPS) Multi Region of Interest (ROI)' analysis module. NNPS refers to the Normalised Noise Power Spectrum.

B.18.4. 'NPS Auto ROI' Module

Additional configuration parameters are required to manage the automatic placement of ROIs within a specified base ROI. These are highlighted in the specification of this module in table B.22.

Inherits from: None			
Inputs	Parameters	Outputs	Effect on Image
Base ROI	**Exclusion ROIs** **ROI Size** **ROI Overlap - X** **ROI Overlap - Y** **Placement Method** **No of ROIs - X** **No of ROIs - Y** Layer Approach Trend Removal Method Normalisation Type Normalisation ROI Data Window Calculate NNPS? Profile Width Axes to Exclude NPS Rebin Interval	**No of ROIs - X** **No of ROIs - Y** Stochastic NPS Image Fixed Pattern NPS Image Intermediate Images 1D NPS Profiles Data Window	None

Table B.22. – Specification of the 'Noise Power Spectrum (NPS) Auto Region of Interest (ROI)' analysis module. NNPS refers to the Normalised Noise Power Spectrum.

Sub regions are positioned within the base ROI in a tile-like fashion. The 'Placement Method' can be set to either automatically fill the base ROI with as many sub-regions as possible, or to place up to a maximum number in the X and Y directions. Regardless of the method chosen, the actual numbers of sub-regions placed in both directions are returned as outputs. When positioning sub-regions they can either be tiled edge to edge, with no overlap, or they can overlap by a quarter or half a sub-region in either direction. Although overlapping sub-regions increases the uncertainty in the NPS calculation for the same number of regions, this can be useful if the image is not large enough to allow placement of sufficient large regions to be able to extract high frequency NPS points with reasonable statistics. Because the automatic placement algorithm may cause sub-regions to be located in parts of the base ROI where may be deliberate structure or artefacts from known problems it is possible to specify a list of ROIs to exclude from the NPS evaluation. These are still placed as normal but are not included in the NPS calculation.

B.18.5. Validation of NPS Algorithm

Validation of the NPS algorithm was performed by using the 'NPS Auto ROI' module to analyse two test images and comparing the results against those generated by OBJ-IQ_reduced and DIMOND3 QA-Distri.

The first image was a synthetic noise image, as suggested by Flynn and Samei[97], and was generated automatically by IQWorks. A 10 layer image was created with each layer having dimensions of 1024×1024 and a pixel size of 0.2 mm $\times$ 0.2 mm. Pixel values were drawn from a Poisson distribution of mean 1000 and according to Flynn and Samei's formalism this image should have a NNPS of 4×10^{-5} mm^2 at all spatial frequencies and in each matrix direction.

A screenshot of the noise image generated by IQWorks is shown in figure B.14 and illustrates the regions automatically placed by the 'NPS Auto ROI' algorithm. The NNPS calculated by the different packages are plotted in figure B.15. It is clear from the figure that all packages perform nearly identically, with the NNPS indeed being constant at the expected value.

As a more difficult test a uniform radiographic flood-field image from one of the Varian OBI systems was analysed. From the results in figure B.16 it is again evident that all packages performed almost identically, even down to revealing structure in the Y direction spectra.

Similar results were also obtained for both images using the more general 'NPS Multi ROI' module, although the configuration of the module was more

time-consuming.

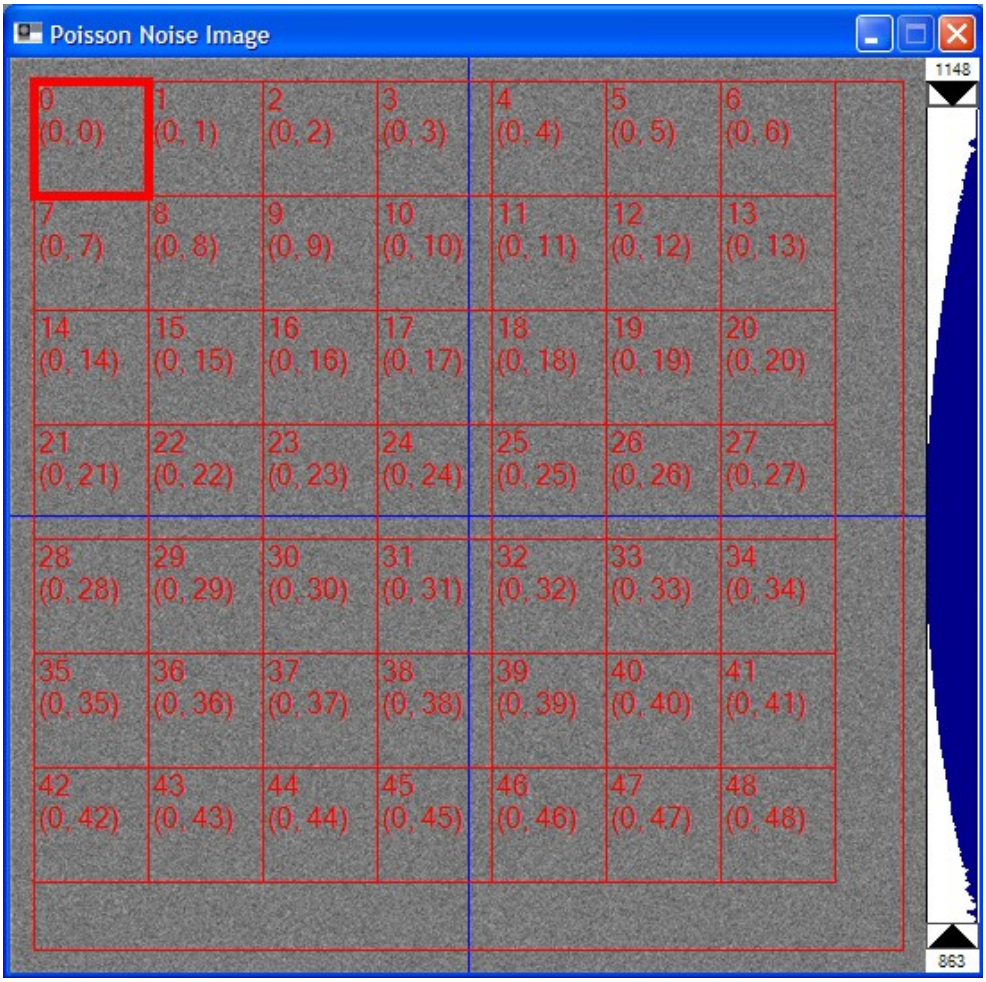


Figure B.14. – Screenshot of the noise image generated by IQWorks to validate the Noise Power Spectrum (NPS) calculation algorithms. Details of the parameters used to generate the image are included in the main text. Regions of interest (ROIs) automatically placed for the analysis are indicated by the red squares, each of which is referenced by a unique number (a, b) where a is the layer number and b is the index of the ROI in the layer. In this example the first layer of a 10 layer image is displayed, so $a = 0$. The bold red square at the top left is the 'reference ROI' against which the results in all other ROIs are normalised.

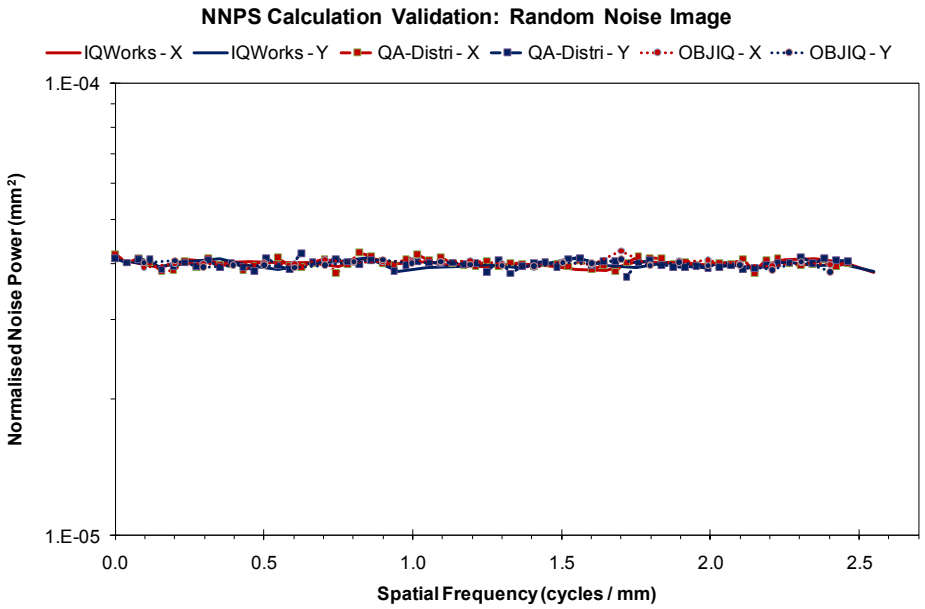


Figure B.15. – Plots of the normalised noise power spectrum (NNPS) in the X and Y directions of the Poisson noise image in figure B.14, as calculated by IQWorks, QA-Distri and OBJ-IQ_reduced.

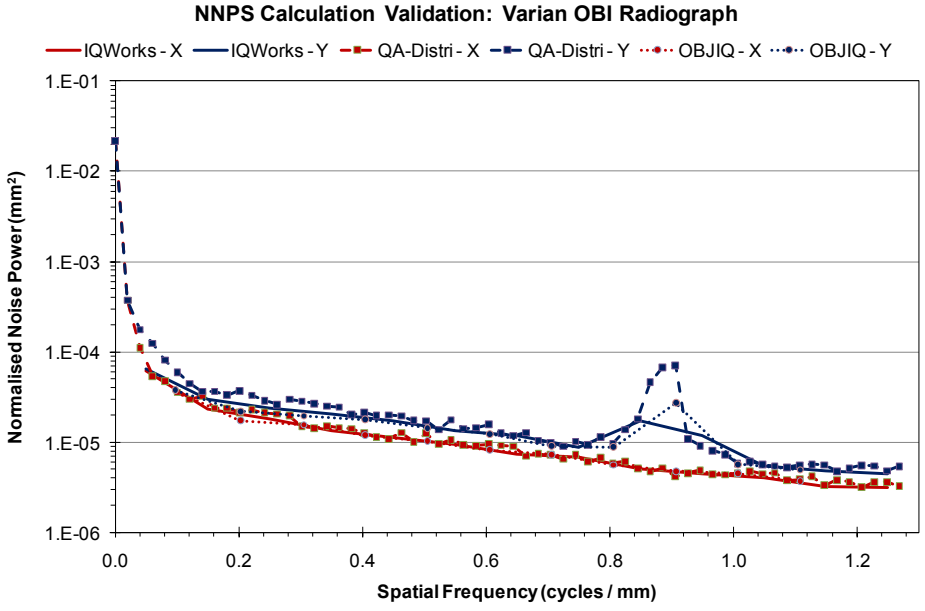


Figure B.16. – Plots of the normalised noise power spectrum (NNPS) in the X and Y directions of a Varian OBI flood radiograph, as calculated by IQWorks, QA-Distri and OBJ-IQ_reduced.

B.19. Detective Quantum Efficiency (DQE)

Inherits from: None			
Inputs	Parameters	Outputs	Effect on Image
MTF Curve NNPS Curve	Number of Quanta Frequency Interval	DQE Curve	None

Table B.23. – Specification of 'Detective Quantum Efficiency (DQE)' analysis module, which accepts modulation transfer function (MTF) and normalised noise power spectrum (NNPS) curves as input.

This module is specified in table B.23 and calculates DQE using equations 2.88 and 2.89. It accepts MTF and NNPS curves output by any of the MTF and NPS analysis modules, along with the number of quanta Q used to form the image. If the value of Q supplied is unity (which is the default) then this module effectively calculates the Noise Equivalent Quanta (NEQ) of the image.

Ideally, MTF and NNPS curves should be provided which are specified at the same frequency interval and over the same frequency range. However, if the curves are defined at different frequency values then both are either interpolated or rebinned to the new interval specified by the 'frequency interval' parameter. In addition, if the curves describe different frequency ranges then that which is defined over the longer range is truncated to match that of the shorter.

B.20. DRR Geometry Checker

Digitally Reconstructed Radiographs (DRRs) are the most widely used reference images against which set-up verification images taken during treatment are compared[284]. Although generated from CT slices by a treatment planning or virtual simulation system they are such an integral component of the radiotherapy process and their quality is influenced by many factors intrinsic to both CT scanning and the particular DRR generation algorithms that they can justifiably be considered a modality in their own right. This analysis module provides a mechanism for quickly verifying the geometrical accuracy of a DRR and can be applied to images of any phantom containing small, easily discernible objects at known locations. However, it is particularly intended for use with a new DRR phantom developed as part of this work and introduced in chapter 5. The specifications of this module are provided in table B.24 and a screenshot of this module being utilised to verify the geometrical accuracy of a DRR for a complex

beam geometry is presented in figure B.17.

Inherits from: None			
Inputs	Parameters	Outputs	Effect on Image
None	Object Coordinates Source-Axis-Distance Source-Image-Distance Gantry Angle Couch Rotation Isocentre Coordinates Imported Treatment Plan	Predicted Object Locations	Overlays Cross-hairs on Objects

Table B.24. – Specification of the analysis module for checking Digitally Reconstructed Radiograph (DRR) geometry.

Lists of physical coordinates of objects in the phantom being imaged are entered into this module. For any combination of isocentre shift, couch rotation, gantry angle, source-axis-distance (SAD) and source-image-distance (SID) the module calculates the expected coordinates of each object on the DRR and overlays cross-hairs at these locations. For example, in the screenshot the DRR is of a phantom containing 4 ball bearings located relative to the scan origin at the physical (x, y, z) coordinates indicated. After calculating the predicted locations of the projections of the ball-bearings in the DRR, given the entered geometrical parameters, the predicted coordinates are displayed numerically alongside the entered values and cross-hairs are overlaid at these positions on the image. By visual inspection, or objectively through simple maxima/minima localisation, it is possible to verify that the predicted and actual locations of the objects match.

Mapping of 3D physical coordinates to a 2D location in the DRR is achieved by modelling the geometry of the linear accelerator. Essentially, the phantom is translated and rotated in 3D space to give points in transformed (x', y') coordinates which can simply be projected onto a virtual imaging plane at the specified SID. For any point in the CT volume $P(x, y, z)$ the location of the projected point on the DRR is given by

$$DRR(X, Y) = M \cdot R_G(\phi) \cdot R_C(\theta) \cdot T(-I) \cdot P(x, y, z) \tag{B.22}$$

where:

M is the simple diverging beam projection matrix,

$R_G(\phi)$ is the transformation due to a gantry rotation ϕ,

$R_C(\theta)$ is the transformation due to a couch rotation θ, and

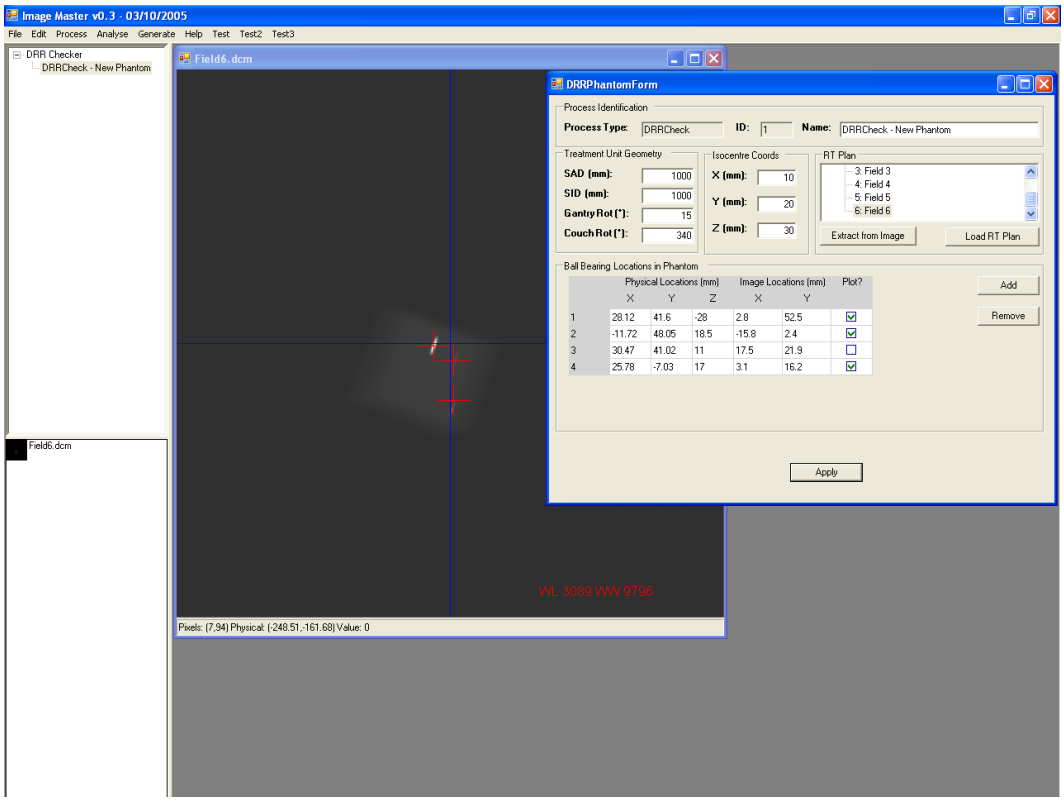


Figure B.17. – Screenshot of IQWorks being used to perform Digitally Reconstructed Radiograph (DRR) geometry checks using the new DRR phantom. Spatial coordinates of ball bearings in the phantom are entered into the dialogue box and a beam geometry is specified. The coordinates of the ball bearings predicted in the DRR image by IQWorks are indicated by the yellow crosses overlayed on the image.

$T(-I)$ is a translation to account for the isocentre $I(x, y, z)$ relative to the origin of coordinates of the CT scan (or the origin of coordinates set in the virtual simulation / treatment planning system).

It is important to note that equation B.22 is non-commutative, that is each transformation must be performed in order: i.e. the isocentre shift must be taken into account first, then the couch rotation, then the gantry rotation and then finally the projection onto the imaging plane. Although it is computationally more efficient to combine the transformation matrices into one single operation it was decided to implement each separately in the IQWorks code because this makes it both far easier to follow what is happening and to extend in the future. Because the calculations are relatively simplistic any performance deficit is negligible.

Entering beam geometry parameters can be tedious and prone to error, especially if a large number of complex geometries are to be considered. A facility has therefore been included to extract these parameters automatically from a DICOM RT Plan object or, if available, from the DICOM metadata of the DRR itself.

Examples of analysing DRRs as part of the clinical optimisation process are included in chapter 5.

B.21. Display Assessment

Display devices are important components of the radiotherapy imaging pathway because they are the final interface between the computer system and the human observer. Because all image processing operations and optimisations which have taken place before the presentation stage may then be altered by the characteristics of the presentation system a badly configured display device has the potential to influence clinical decisions. Image presentation science is a discipline in its own right and an in depth treatment is beyond the scope of this work. However, an analysis module for the assessment, calibration and optimisation of displays in line with DICOM[374] and AAPM TG18[305] guidance has been included in IQWorks so that the framework described in this thesis can be applied at each stage of radiotherapy imaging, from image acquisition through to presentation to the observer by a display device.

Performance assessment of displays is similar to that of imaging modalities and includes characterisation of spatial resolution (MTF), noise (NNPS), geometric linearity and signal transfer properties in terms of the light output

(luminance) as a function of digital driving level (DDL). MTF and NNPS can be calculated through taking photographs of standard test patterns using a scientific grade digital camera then processing these using the appropriate IQWorks analysis modules as for other imaging modalities, albeit taking into account the characteristics of the digital camera. Measurement of signal transfer properties is performed using a photometer. This is traditionally a lengthy process involving displaying a series of test patterns with patches at specified DDLs and measuring the luminance of these.

Being a complex, non-linear system, the human eye responds differently depending upon the absolute power density of light falling on the retina. A model developed by Barten[16, 17] divides the luminance range observable by the human visual system into 'just-noticeable-differences' (JNDs) which correspond to changes in luminance at each luminance point which an average observer would just be able to discern. A linearised display device is one which conforms to this Barten model, which is encapsulated by the DICOM Grayscale Standard Display Function (GSDF), shown in figure B.18[374]. If a device's native curve deviates from the GSDF then a look-up table (LUT) can be calculated which is applied to the DDLs before they are sent to the display driving hardware.

Assessment of display linearity is performed by IQWorks 'Display Assessment' module. This interfaces with photometric instruments to directly measure and then record luminance (or illuminance for projectors) resulting from test patterns displayed at specified DDL operating points. To facilitate rapid display characterisation the module is capable of automatically stepping through a complete range of operating points and measuring the photometric output at each. It then compares these against the GSDF using the methodologies described in the DICOM standard[374] and AAPM TG 18 report[305], presents the results in graphically, and calculates the LUT required to 'calibrate' the display against the GSDF. A further measurement cycle can then be performed with the LUT active so that the successfulness of the calibration can be evaluated. TG 18 specify tolerance bands around the GSDF for display devices being used for different clinical applications and the analysis module indicates whether the performance of a given devices falls within these. Once a LUT has been calculated and accepted it can be applied routinely at login time by using a companion software package 'Auto LUT', developed alongside IQWorks as part of this work[297].

Interfacing with measurement instruments is achieved by inheriting from

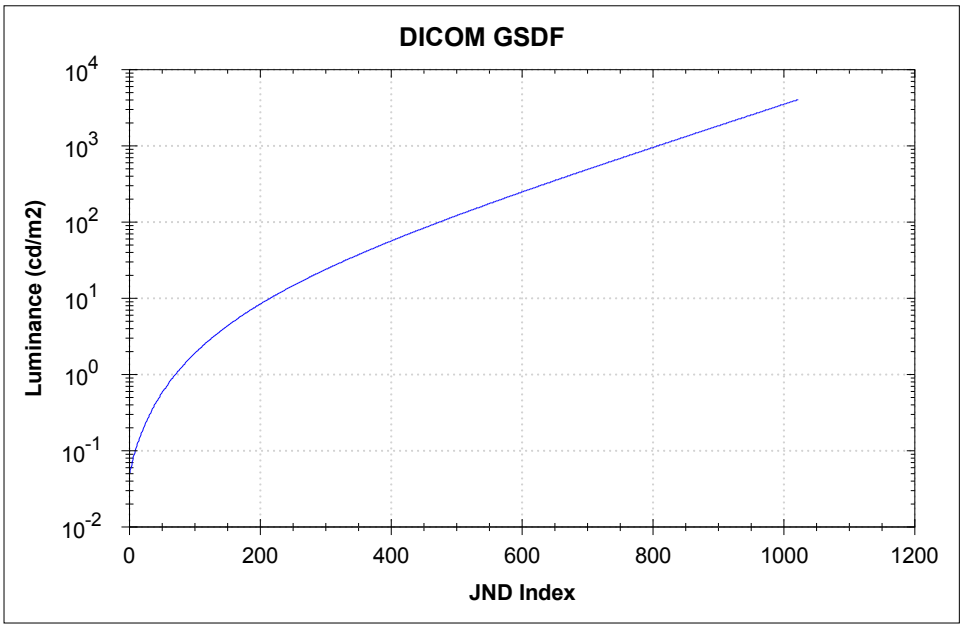


Figure B.18. – DICOM Grayscale Standard Display Function (GSDF). Luminance is plotted as a function of the Barten 'Just Noticeable Distance' (JND) index.

an abstract base class and is relatively straightforward using manufacturers' programming interfaces. To date, interfaces have been written to communicate with the IBA LXPlus, LXCan and LXChroma photometers, the X-Rite eye-one Display 2 / LT photometric pod and the Macam L202 photometer. It is hoped third parties will write further interfaces to accommodate other instruments in the future.

It is also important to consider ambient light levels in work areas where digital images are being displayed and interpreted because any additional light shifts the display's curve up the GSDF so that small changes in light output are less discernible by the human visual system. A datalogging facility has been included in this module which takes periodic readings of environmental light levels and dumps the results to an Excel-readable CSV file. This works with any of the interfaced photometric instruments.

Examples of the 'Display Assessment' module being applied to radiotherapy applications are provided in chapter 5.

B.22. Conclusion

IQWorks includes a wide range of fundamental calculation modules which can be combined in an analysis tree to perform complex and automated processing of images. Although analysis modules have been implemented covering the majority of objective image quality metrics introduced in chapter 2 there will inevitably be situations when new algorithms are required. Built on an extensible, object-orientated framework it is hoped significant new functionality can be added without users requiring to construct new modules from scratch.

Appendix C.

Metrics Calculated by IQWorks

C.1. Large-Area (Macroscopic or Spatial Domain) Metrics

1. Using regions of interest:
    a) Basic ROI statistics (mean, standard deviation, variance, max, min, etc.)
    b) Contrast
    c) Signal-to-noise ratio (SNR)
    d) Difference or detail SNR (dSNR) / contrast-to-noise ratio (CNR)
    e) Coefficient of Variation (CoV)
    f) Variance map
    g) Gradient analysis
2. By extracting line or area profiles:
    a) Coefficient of Variation (CoV)
    b) Integral Uniformity ($U_{i(+)}$ and $U_{i(-)}$)
    c) Differential Uniformity (U_d)

C.2. Spatial-Frequency Domain Metrics

1. Noise Power Spectrum (NPS)
2. Normalised Noise Power Spectrum (NNPS)
3. Modulation Transfer Function (MTF)

   a) From an impulse (i.e. point) object – within an image plane or down a stack of images
   b) From an edge
   c) From a line
   d) From a random field

4. Noise Equivalent Quanta (NEQ)
5. Detective Quantum Efficiency (DQE)
6. 'Reduced', single numerical value descriptors of the above

C.3. Geometrical Factors

1. In a 2D image plane:
   a) Object dimensions (based on detected edge)
   b) Pixel size in either matrix direction
   c) Distances and angles between specified (auto-detected) points
2. In a stack of image, using an appropriate test object:
   a) Slice thickness
   b) Slice location

Appendix D.

Papers and Conference Presentations

D.1. Possible Publications

- A uniform software framework for the objective assessment and optimisation of image quality.
  - Cover applicability to CT, IEC DQE evaluation for digital radiography and MRI.
- A uniform framework for the objective assessment and optimisation of radiotherapy image quality.
  - Covering general applicability to radiotherapy modalities. Emphasise use in commissioning cone-beam CT.
  - Present result that care needs to be exercised when determining HU / CT no calibration curve.
- Digital radiography: application of new image quality evaluation techniques to old phantoms.
- DRR Phantom. Covering optimisation techniques and comparing different platforms.
- Application of framework to MRI. (Very little on general MR QA in the literature.)
- EPID performance evaluation and optimisation: A fresh approach
- A new approach to DICOM calibration of digital display devices.

D.2. Citable Conference Papers

Convert to standard citation format.

- Reilly AJ, MacLeod AS, Initial Experience in Commissioning a Dedicated Oncology CT Scanner. Proceedings of the UK Radiological Congress. 2002, 103. ISBN 0-905749-49-9
- Reilly AJ, Erridge SC, Wilson D, Warnock J, Goodall L, Peoples S, Lord H, Robinson A, Thwaites DI. Assessment of Patient Set-up Errors: Who Should Measure It and With What Technique. Clinical Oncology Supplement 1. 2005, 17, S26-S27
- Reilly AJ, Skrzyński W, Thwaites DI. EPID Quality Assurance: A Fresh Approach. Clinical Oncology Supplement 1. 2005, 17, S19
- Reilly AJ, MacLeod AS, Thwaites DI. A Phantom for Checking and Optimising Digitally Reconstructed Radiographs. Proceedings of the UK Radiological Congress. 2005
- Reilly AJ. Investigation of a Photographic Grade Digital Camera for the Evaluation of Image Display Devices. Proceedings of the UK Radiological Congress. 2005
- Reilly AJ. Automated Image Analysis software for Radiotherapy and Diagnostic CT Quality Assurance. Proceedings of the UK Radiological Congress. 2005
- Skrzyński W, Reilly AJ, Thwaites DI, Bulski W. EPID Image Quality: Las Vegas Phantom as an Objective Tool. Proceedings of the 13th Congress of the Polish Society of Medical Physics. 2005
- Reilly AJ, Skrzyński W, McKerracher, C, Thwaites DI. A New Performance Phantoms for EPID Quality Assurance. Proceedings of the 9th International Workshop on Electronic Portal Imaging (EPI2K6). 2006
- Reilly AJ, Erridge SC, Ironside J, Little F, Junor E, McLaren D, Price A, Wilson D. A novel tool for assessing set-up uncertainties applied to radical radiotherapy patients. Clinical Oncology. 2007
- Reilly AJ, MacLeod AS, Thwaites DI. Digitally Reconstructed Radiographs: a new phantom for verifying accuracy and optimising quality. Clinical Oncology. 2007

- Reilly AJ, Thwaites DI. Automated Analysis Software for the Objective Assessment and Optimisation of Radiotherapy Image Quality. International Journal of Radiation Oncology*Biology*Physics. 2007.
- Reilly AJ, Thwaites DI. IQ Works: An Automated Image Analysis Framework for the Objective Assessment and Optimisation of Image Quality. Radiology. 2007
- Reilly AJ, Thwaites DI. A new phantom for the objective performance evaluation of Electronic Portal Imaging Devices (EPIDs). Radiology. 2007.
- Reilly AJ, Weir N. A new package and the evaluation of detectors for display and projector luminance response curve calibration against the DICOM Part 14 Grayscale Standard Display Function (GSDF). Radiology. 2007.

D.3. Other Papers

- Reilly AJ, Erridge S. A Tool for Calculating and Tracking Patient Set-up Uncertainties. IPEM Biennial Radiotherapy meeting. 2006
- Reilly AJ, Skrzyński W, McKerracher C, Thwaites DI. A New Performance Phantom for EPID Quality Assurance. IPEM Biennial Radiotherapy Meeting. 2006
- Reilly AJ, Macleod AS, Thwaites DI. A Phantom for Checking and Optimising Digitally Reconstructed Radiographs. IPEM Biennial Radiotherapy Meeting. 2006
- Reilly AJ. A Generic Software Framework for Imaging Quality Assurance. 8th CT User Group Meeting. 2006
- Reilly AJ. Automated image analysis software for the objective assessment and optimisation of radiotherapy image quality. Scottish+ Radiotherapy Physics Meeting. 2007
- Reilly AJ and Erridge SC. A novel tool for calculating and tracking patient set-up uncertainties. Scottish+ Radiotherapy Physics Meeting. 2007
- Reilly AJ. IQ Works: Automated Image Quality Analysis. IPEM Meeting on Software Image Analysis for QA. 2007.

- Reilly AJ, Sankar A, MacLeod A. MapCheck for IMRT Verification and EPID Dosimetry. Invited review at Sun Nuclear MapCheck Symposium, ESTRO, Barcelona. 2007.
- Reilly AJ, Weir N. IQWorks: New Software for the Calibration and Quality Assurance of Displays and Projectors. IPEM meeting on Imaging Quality Assurance. 2007.
- Reilly AJ. IQWorks: New Software for the Calibration and QA of Radiotherapy Displays. Scottish+ Radiotherapy Physics Meeting. 2008.
- Reilly AJ. Using IQWorks for Automated Image Quality Evaluation in Radiotherapy and Diagnostic Imaging. Invited teaching session at IPEM Annual Scientific Meeting. 2008.
- Reilly AJ. IQWorks Update. CT User Group Meeting. 2008.
- Reilly AJ. Determining CT and CBCT Hounsfield Unit calibration curves for treatment planning – some unexpected results. Radiotherapy Imaging User Group Meeting. 2008

Appendix E.

Modalities to which IQWorks has been Applied

IQWorks has been successfully used to process images from the following modalities. (Only those encountered in radiotherapy are discussed in the main body of the thesis.)

1. **CT - Computed Tomography**
    a) Conventional CT scanner
    b) CT simulator
    c) Cone-Beam CT - radiotherapy simulation
    d) Cone-Beam CT - radiotherapy verification (systems integrated with linac gantry)
    e) Cone-Beam CT - diagnostic systems (general and dental)
    f) Tomotherapy
2. **DRR - Digitally Reconstructed Radiograph**
    a) GE Advantage Sim
    b) Varian Eclipse
3. **Kilovoltage X-Ray Projection Imaging**
    a) RT Simulator- image intensifier based
    b) RT Simulator - aSi flat panel based
    c) RT Verification Imaging - aSi flat panel detector mounted on linac gantry
    d) CR plates - diagnostic imaging

   e) CR plates - radiotherapy imaging
   f) Direct digital diagnostic imaging
   g) Direct digital mammography

4. EPI - Electronic Portal Imaging
   a) Varian aS500, aS500-II and aS1000
   b) Elekta iView GT

5. MRI - Magnetic Resonance Imaging
   a) Various diagnostic systems

6. Soft Copy Display Devices
   a) CRT devices
   b) LCD flat panel devices
   c) Projector systems

Appendix F.

Computer Software

The complete IQWorks software package, including source code, a compiled executable and example images, can be downloaded from the IQWorks website: `http://www.iqworks.org`.

F.1. IQWorks Licence Agreement

F.1.1. IQWorks Licensing

IQWorks consists of a number of components, in addition to the core IQWorks system itself. Each of these components has been released, and is distributed with IQWorks, under its own licence.

The core IQWorks licence agreement is included below. Details of agreements pertaining to the 3rd party components utilised by IQWorks are included on the website at `http://www.iqworks.org/licence`.

F.1.2. IQWorks Core System v0.6

The IQWorks Core System is distributed under version 3 of the GNU General Public Licence (GPL), with the following additional conditions:

1. This software should be used only by qualified medical imaging professionals who understand the algorithms being implemented and the results presented. It must not be used directly to make decisions affecting the clinical management of patients. Being an early release, there will certainly be bugs, and no guarantee can be given for the overall stability

of the package. Therefore, it should not be installed on any mission critical computer system. Furthermore, it must not be installed on systems classified as "medical devices" by any commonly accepted definition of the term.

2. Neither the names of the names of the original developers (Andrew J Reilly, NHS Lothian, The University of Edinburgh or Oxford Radcliffe Hospitals NHS Trust) nor the names of other contributors may be used to endorse or promote products derived from this software without specific prior written permission.

Most third party components used by and bundled with IQWorks are implemented as DLLs or ActiveX controls. Their source code is not generally incorporated into IQWorks and, in most instances, IQWorks will run substantially without them.

References

[1] Abramoff MD, Magelhaes PJ, Ram SJ. Image Processing with ImageJ. Biophotonics International. 2004;11(7):36–42.

[2] ACR. ACR CT Accreditation Phantom: Instructions for Used;.

[3] Adler, Murphy, Chang, Hancock. Image-guided robotic radiosurgery. Neurosurgery. 1999 Jun;44(6):1299–306; discussion 1306–7.

[4] Ahnesjo A, Aspradakis MM. Dose calculations for external photon beams in radiotherapy. Physics in Medicine and Biology. 1999;44(11):R99–R155. Available from: http://stacks.iop.org/0031-9155/44/R99.

[5] Aird EGA. Second cancer risk, concomitant exposures and IRMER(2000). Br J Radiol. 2004 Dec;77(924):983–985.

[6] Aird EGA, Conway J. CT simulation for radiotherapy treatment planning. Br J Radiol. 2002;75(900):937–949. Available from: http://bjr.birjournals.org/cgi/content/abstract/75/900/937.

[7] Alaei P, Ding G, Guan H. Inclusion of the dose from kilovoltage cone beam CT in the radiation therapy treatment plans. Med Phys. 2010 Jan;37(1):244–248.

[8] Amer A, Marchant T, Sykes J, Czajka J, Moore C. Imaging doses from the Elekta Synergy X-ray cone beam CT system. Br J Radiol. 2007 Jun;80(954):476–482. Available from: http://dx.doi.org/10.1259/bjr/80446730.

[9] Andreo P. Uncertainties in dosimetric data and beam calibration. Int J Radiat Oncol Biol Phys. 1990 Nov;19(5):1233–1247.

[10] Antypas C, Pantelis E. Performance evaluation of a CyberKnife G4 image-guided robotic stereotactic radiosurgery system. Physics in Medicine and Biology. 2008;53(17):4697–4718. Available from: http://stacks.iop.org/0031-9155/53/4697.

[11] Aquilab. ARTISCAN(TM) Brochure; 2009. Available from: http://www.aquilab.com/images/stories/Artiscan/artiscan-gb.pdf.

[12] Atwood B. Area Measurement: Planimeters & Green's Theorem; 2010. Available from: http://www.attewode.com/Calculus/AreaMeasurement/area.htm.

[13] Aufrichtig R. Comparison of low contrast detectability between a digital amorphous silicon and a screen-film based imaging system for thoracic radiography. Med Phys. 1999 Jul;26(7):1349–1358.

[14] Barrett H. Object assessment of image quality - effects of quantum noise and object variability [Article]. Journal of the Optical Society of America A-Optics Image Science and Vision. 1990 JUL;7(7):1266–1278.

[15] Barrett H, Smith W, KJ M, TD M, Fiete R. Quantifying the performance of imaging-systems [Article]. Proceedings of the Society of Photo-Optical Instrumentation Engineers. 1985;535:65–69.

[16] Barten PGJ. Spatiotemporal model for the contrast sensitivity of the human eye and its temporal aspects. Human Vision, Visual Processing, and Digital Display IV. 1993;1913(1):2–14. Available from: http://link.aip.org/link/?PSI/1913/2/1.

[17] Barten PGJ. Physical model for the contrast sensitivity of the human eye. vol. 1666. Rogowitz BE, editor. SPIE; 1992. Available from: http://link.aip.org/link/?PSI/1666/57/1.

[18] Battista JJ, Bronskill MJ. Compton scatter imaging of transverse sections: an overall appraisal and evaluation for radiotherapy planning. Phys Med Biol. 1981 Jan;26(1):81–99.

[19] Baumann M, Krause M, Zips D, Petersen C, Dittmann K, Dörr W, et al. Molecular targeting in radiotherapy of lung cancer. Lung Cancer. 2004 Aug;45 Suppl 2:S187–S197. Available from: http://dx.doi.org/10.1016/j.lungcan.2004.07.975.

[20] Bedford JL, Warrington AP. Commissioning of Volumetric Modulated Arc Therapy (VMAT). International Journal of Radiation Oncology*Biology*Physics. 2009;73(2):537 – 545. Available from: http://www.sciencedirect.com/science/article/B6T7X-4VC63P7-7/2/81efd40c9222fc6007cd3b61d6229ff9.

[21] Benitez RB, Ning R, Conover D, Liu S. Measurements of the Modulation Transfer Function, Normalized Noise Power Spectrum and Detective Quantum Efficiency for two flat panel detectors: A Fluoroscopic and a Cone Beam Computer Tomography Flat Panel Detectors. Journal of X-Ray Science and Technology. 2009;17:279–293.

[22] Bentzen SM. Theragnostic imaging for radiation oncology: dose-painting by numbers. The Lancet Oncology. 2005;6(2):112 – 117. Available from: http://www.sciencedirect.com/science/article/B6W85-4FC34VY-S/2/15cf4f0520037f634af9fd4c162656a6.

[23] Bentzen SM, Heeren G, Cottier B, Slotman B, Glimelius B, Lievens Y, et al. Towards evidence-based guidelines for radiotherapy infrastructure and staffing needs in Europe: the ESTRO QUARTS project. Radiotherapy and Oncology. 2005;75(3):355 – 365. Available from: http://www.sciencedirect.com/science/article/B6TBY-4FPX2FT-1/2/0067e418e37fb6f0599440e1458eccf6.

[24] Berger L, François P, Gaboriaud G, Rosenwald JC. Performance optimization of the Varian aS500 EPID system. J Appl Clin Med Phys. 2006;7(1):105–114.

[25] Berrang TS, Truong PT, Popescu C, Drever L, Kader HA, Hilts ML, et al. 3D Ultrasound Can Contribute to Planning CT to Define the Target for Partial Breast Radiotherapy. International Journal of Radiation Oncology*Biology*Physics. 2009;73(2):375 – 383. Available from: http://www.sciencedirect.com/science/article/B6T7X-4T5D9MR-5/2/90a45da218d8be5eb73b7f4bd082ae55.

[26] Biomedical F. Nuclear Associates 76-410-4130 and 76-411 AAPM CT Performance Phantom Users Manual. Cleveland, Ohio; 2005.

[27] BIR. Central Axis Depth Dose Data for Use in Radiotherapy: BJR Supplement 25. London: British Institute of Radiology; 1996.

[28] BIR. Central Axis Depth Dose Data for Use in Radiotherapy: BJR Supplement 17. London: British Institute of Radiology; 1983.

[29] Boas ML. Mathematical Methods in the Physical Sciences, 2nd Edition. 2nd ed. New York: John Wiley and Sons; 1983.

[30] Boedeker KL, Cooper VN, McNitt-Gray MF. Application of the noise power spectrum in modern diagnostic MDCT: part I. Measurement of noise power spectra and noise equivalent quanta. Physics in Medicine and Biology. 2007;52(14):4027–4046. Available from: http://stacks.iop.org/0031-9155/52/4027.

[31] Boedeker KL, McNitt-Gray MF. Application of the noise power spectrum in modern diagnostic MDCT: part II. Noise power spectra and signal to noise. Physics in Medicine and Biology. 2007;52(14):4047–4061. Available from: http://stacks.iop.org/0031-9155/52/4047.

[32] Bondiau PY, Malandain G, Chanalet S, Marcy PY, Habrand JL, Fauchon F, et al. Atlas-based automatic segmentation of MR images: Validation study on the brainstem in radiotherapy context. International Journal of Radiation Oncology*Biology*Physics. 2005;61(1):289 – 298. Available from: http://www.sciencedirect.com/science/article/B6T7X-4F4H43H-1C/2/cad006f5304d696b5b4085d06c1cac59.

[33] Boone JM. Determination of the presampled MTF in computed tomography. Medical Physics. 2001;28(3):356–360. Available from: http://link.aip.org/link/?MPH/28/356/1.

[34] Boone JM, Cody DD, Fisher JR, Frey GD, Glasser H, Gray JE, et al. AAPM Report 74: Quality Control in Diagnostic Radiology. AAPM; 2002.

[35] Brada M, Pijls-Johannesma M, Ruysscher DD. Proton therapy in clinical practice: current clinical evidence. J Clin Oncol. 2007 Mar;25(8):965–970. Available from: http://dx.doi.org/10.1200/JCO.2006.10.0131.

[36] Brahme A. Dosimetric precision requirements in radiation therapy. Acta Radiol Oncol. 1984;23(5):379–391.

[37] Brennan PC, McDonnell S, O'Leary D. Increasing film-focus distance (FFD) reduces radiation dose for x-ray examinations. Radiat Prot Dosimetry. 2004;108(3):263–268. Available from: http://dx.doi.org/10.1093/rpd/nch029.

[38] Brenner DJ. Effective dose: a flawed concept that could and should be replaced. Br J Radiol. 2008 Jul;81(967):521–523. Available from: http://dx.doi.org/10.1259/bjr/22942198.

[39] Brlek S, Labelle G, Lacasse A. The discrete Green Theorem and some applications in discrete geometry. Theoretical Computer Science. 2005;346(2-3):200 – 225. In memoriam: Alberto Del Lungo (1965-2003). Available from: http://www.sciencedirect.com/science/article/B6V1G-4H2PHW8-1/2/38c35d6bdc31c2edb2f15aad593843c2.

[40] BSI. BS EN 12588:2006 Lead and lead alloys – Rolled lead sheet for building purposes; 2006.

[41] BSI. BS EN 60601-1:2006. Medical electrical equipment. General requirements for basic safety and essential performance; 2006.

[42] Buckey C. Characterization of a Varian aS1000 EPID; 2010. WWW. Available from: http://ric.uthscsa.edu/personalpages/lancaster/DI2_Projects_2010/EPID_characterization.pdf.

[43] Buckle AH. Varying MLC end projection size: an effect of the double-focused Siemens multileaf collimator. Physics in Medicine and Biology. 2006;51(8):N147–N155. Available from: http://stacks.iop.org/0031-9155/51/N147.

[44] Burgess AE. The Rose model, revisited. J Opt Soc Am A. 1999;16(3):633–646. Available from: http://josaa.osa.org/abstract.cfm?URI=josaa-16-3-633.

[45] Carden J, Keat N, Lewis M, Edyvean S. IGE HiSpeed FX/i CT Scanner Technical Evaluation. Medical Devices Agency (ImPACT); 1999. MDA Evaluation Report MDA/99/65.

[46] Carlotto MJ. Histogram Analysis Using a Scale-Space Approach. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1987;PAMI-9(1):121–129.

[47] CEC. EUR 16260 EN: European Guidelines on Quality Criteria for Diagnostic Radiographic Images. Carmichael JHE, Maccia C, Moores BM, Oestmann JW, Schibilla H, Teunen D, et al., editors. European Commission; 1995.

[48] Chang SD, Adler JR. Robotics and radiosurgery–the cyberknife. Stereotact Funct Neurosurg. 2001;76(3-4):204–208.

[49] Chen L, Nguyen TB, Élan Jones, Chen Z, Luo W, Wang L, et al. Magnetic Resonance-Based Treatment Planning for Prostate Intensity-Modulated Radiotherapy: Creation of Digitally Reconstructed Radiographs. International Journal of Radiation Oncology*Biology*Physics. 2007;68(3):903 – 911. Available from: http://www.sciencedirect.com/science/article/B6T7X-4NVRVV1-5/2/3595bc36cd2d1c3161bb2880f2d01517.

[50] Chesters MS, Hay GA. Quantitative relation between detectability and noise power. Physics in Medicine and Biology. 1983;28(10):1113–1125. Available from: http://stacks.iop.org/0031-9155/28/1113.

[51] Cigna AA, Nassisi D, Masenga D, Raffo R, Rotta P. Dose due to scattered radiation in external radiotherapy: a prostate cancer case history. Radiat Prot Dosimetry. 2004;108(1):27–32.

[52] Clarke R, ICRP. 21st century challenges in radiation protection and shielding: draft 2005 recommendations of ICRP. Radiat Prot Dosimetry. 2005;115(1-4):10–15.

[53] Cochran J; 2010. Available from: http://www.developingfor.net/free-code/polygonf.

[54] Coltman JW. The Specification of Imaging Properties by Response to a Sine Wave Input. J Opt Soc Am. 1954;44(6):468–469. Available from: http://www.opticsinfobase.org/abstract.cfm?URI=josa-44-6-468.

[55] Constantinou C, Harrington JC, DeWerd LA. An electron density calibration phantom for CT-based treatment planning computers. Med Phys. 1992;19(2):325–327.

[56] Court L, Rosen I, Mohan R, Dong L. Evaluation of mechanical precision and alignment uncertainties for an integrated CT/LINAC system. Med Phys. 2003 Jun;30(6):1198–1210.

[57] Cowen AR, Launders JH, Jadav M, Brettle DS. Visibility of microcalcifications in computed and screen-film mammography. Physics in Medicine and Biology. 1997;42(8):1533–1548. Available from: http://stacks.iop.org/0031-9155/42/1533.

[58] Cranley K, Gilmore BJ, Fogarty GWA, Desponds L, Sutton D. IPEM Report 78: Catalogue of Diagnostic X-ray Spectra and Other Data. IPEM; 1997.

[59] Crawley MJ. Statistics – An Introduction Using R. Chichester, England: John Wiley and Sons; 2005.

[60] Cremers F, Frenzel T, Kausch C, Albers D, Schönborn T, Schmidt R. Performance of electronic portal imaging devices (EPIDs) used in radiotherapy: Image quality and dose measurements. Medical Physics. 2004;31(5):985–996. Available from: `http://link.aip.org/link/?MPH/31/985/1`.

[61] CRUK. CancerStats Key Facts: All Cancers; 2008. Available from: `http://publications.cancerresearchuk.org/WebRoot/crukstoredb/CRUK_PDFs/AllCancersKeyFactsJune08.pdf`.

[62] Cunningham IA. 2: Applied Linear-Systems Theory. In: Jacob Beutel RLV Harold L Kundel, editor. Handbook of Medical Imaging. vol. 1. Physics and Psychophysics. 1st ed. SPIE Publications; 2000. p. 79–159.

[63] Cunningham IA, Reid BK. Signal and noise in modulation transfer function determinations using the slit, wire, and edge techniques. Med Phys. 1992;19(4):1037–1044.

[64] Dawson P. Patient dose in multislice CT: why is it increasing and does it matter? Br J Radiol. 2004;77 Spec No 1:S10–S13.

[65] Dearnaley DP, Hall E, Lawrence D, Huddart RA, Eeles R, Nutting CM, et al. Phase III pilot study of dose escalation using conformal radiotherapy in prostate cancer: PSA control and side effects. Br J Cancer. 2005 Feb;92(3):488–498. Available from: `http://dx.doi.org/10.1038/sj.bjc.6602301`.

[66] Delaney G, Jacob S, Featherstone C, Barton M. The role of radiotherapy in cancer treatment: estimating optimal utilization from a review of evidence-based clinical guidelines. Cancer. 2005 Sep;104(6):1129–1137. Available from: `http://dx.doi.org/10.1002/cncr.21324`.

[67] Dendy P, Heaton B. Physics for Diagnostic Radiology. 2nd ed. Institute of Physics; 1999.

[68] Diallo I, Lamon A, Shamsaldin A, Grimaud E, de Vathaire F, Chavaudra J. Estimation of the radiation dose delivered to any point outside the target volume per patient treated with external beam radiotherapy. Radiother Oncol. 1996 Mar;38(3):269–271.

[69] Dietze G, Harrison JD, Menzel HG. Effective dose: a flawed concept that could and should be replaced. Comments on a paper by D J Brenner (Br J Radiol 2008;81:521-3). Br J Radiol. 2009 Apr;82(976):348–50; author reply 350–1. Available from: http://dx.doi.org/10.1259/bjr/91937653.

[70] Ding G. WE-A-203-01: X-Ray Imaging Dose from IGRT. vol. 37. AAPM; 2010. p. 3409–3409. Available from: http://link.aip.org/link/?MPH/37/3409/1.

[71] Ding GX, Munro P, Pawlowski J, Malcolm A, Coffey CW. Reducing radiation exposure to patients from kV-CBCT imaging. Radiother Oncol. 2010 Sep;Available from: http://dx.doi.org/10.1016/j.radonc.2010.08.005.

[72] Ding GX, Pawlowski JM, Coffey CW. A correction-based dose calculation algorithm for kilovoltage x rays. Med Phys. 2008 Dec;35(12):5312–5316.

[73] Dobbins JT. 1: Medical Physics and Psychophysics. In: Jacob Beutel RLV Harold L Kundel, editor. Handbook of Medical Imaging. vol. 1. Physics and Psychophysics. 1st ed. SPIE Publications; 2000. .

[74] Dobbins JT. Effects of undersampling on the proper interpretation of modulation transfer function, noise power spectra, and noise equivalent quanta of digital imaging systems. Med Phys. 1995 Feb;22(2):171–181.

[75] Dobbins JT, Ergun DL, Rutz L, Hinshaw DA, Blume H, Clark DC. Dqe(F) of 4 Generations of Computed Radiography Acquisition Devices. Medical Physics. 1995;22(10):1581–1593. Available from: ISI:A1995TA81700003.

[76] Dobbins JT, Samei E, Ranger NT, Chen Y. Intercomparison of methods for image quality characterization. II. Noise power spectrum. Medical Physics. 2006 May;33(5):1466–1475.

[77] Dong L, Boyer AL. An objective method for evaluating electronic portal imaging devices. Medical Physics. 1994;21(6):755–760. Available from: http://link.aip.org/link/?MPH/21/755/1.

[78] Doyle P. Assessment and Optimisation of Digital Radiography Systems for Clinical Use. University of Glasgow; 2008.

[79] Droege RT. A practical method to routinely monitor resolution in digital images. Medical Physics. 1983;10(3):337–343. Available from: http://link.aip.org/link/?MPH/10/337/1.

[80] Droege RT, Morin RL. A practical method to measure the MTF of CT scanners. Medical Physics. 1982;9(5):758–760. Available from: http://link.aip.org/link/?MPH/9/758/1.

[81] Dutreix A. When and how can we improve precision in radiotherapy? Radiother Oncol. 1984 Dec;2(4):275–292.

[82] Dörr W, Herrmann T. Cancer induction by radiotherapy: dose dependence and spatial relationship to irradiated volume. J Radiol Prot. 2002 Sep;22(3A):A117–A121.

[83] EC. Council Directive 97/43/EURATOM; 1997.

[84] Edyvean S. ImPACT Course Lecture 9: Understanding Image Quality and Dose. ImPACT; 2007. Available from: http://www.impactscan.org/slides/course07/lect9/frame.htm.

[85] Elbalaa ZAK, Foulquier JN, balaa HE, Orthuon A, Touboul E. Quality assurance of enhanced dynamic wedge using the aS500-II, EPID. Cancer Radiother. 2009 Dec;13(8):731–739. Available from: http://dx.doi.org/10.1016/j.canrad.2009.08.142.

[86] Elbalaa ZAK, Foulquier JN, Orthuon A, Elbalaa H, Touboul E. Role of "the frame cycle time" in portal dose imaging using an aS500-II EPID. Phys Med. 2009 Sep;25(3):148–153. Available from: http://dx.doi.org/10.1016/j.ejmp.2009.01.004.

[87] Elekta. Precise Treatment System; 2008. Available from: http://www.elekta.com/assets/precision_radiation_therapy/0185%2010-08%20PTSystem%20brochure.pdf.

[88] Elekta. Precision Multileaf Collimator; 2005. Available from: http://www.elekta.com/assets/precision_radiation_therapy/0188%2012-06%20MLCi.pdf.

[89] van Elmpt W, McDermott L, Nijsten S, Wendling M, Lambin P, Mijnheer B. A literature review of electronic portal imaging for radiotherapy dosimetry. Radiother Oncol. 2008 Sep;88(3):289–309. Available from: http://dx.doi.org/10.1016/j.radonc.2008.07.008.

[90] van Elmpt W, Nijsten S, Petit S, Mijnheer B, Lambin P, Dekker A. 3D In Vivo Dosimetry Using Megavoltage Cone-Beam CT

and EPID Dosimetry. International Journal of Radiation Oncology*Biology*Physics. 2009;73(5):1580 – 1587. Available from: http://www.sciencedirect.com/science/article/B6T7X-4VWCT67-D/2/8253255403fc81b4ac4e16e9cbc7e156.

[91] Erridge SC, Featherstone C, Chalmers R, Campbell J, Stockton D, Black R. What will be the radiotherapy machine capacity required for optimal delivery of radiotherapy in Scotland in 2015? Eur J Cancer. 2007 Aug;43(12):1802–1809. Available from: http://dx.doi.org/10.1016/j.ejca.2007.05.022.

[92] Esch AV, Depuydt T, Huyskens DP. The use of an aSi-based EPID for routine absolute dosimetric pre-treatment verification of dynamic IMRT fields. Radiotherapy and Oncology. 2004;71(2):223 – 234. Available from: http://www.sciencedirect.com/science/article/B6TBY-4C4C4VR-5/2/dbbacadeac1f690df9cf77264db9573c.

[93] EU. Sentinel: Safety and Efficacy for New Techniques and Imaging Using New Equipment to Support European Legislation. Euratom Directorate General for Research; 2007.

[94] Fabbri S, Taibi A, Longo R, Marziani M, Olivo A, Pani S, et al. Signal-to-noise ratio evaluation in dual-energy radiography with synchrotron radiation. Phys Med Biol. 2002 Nov;47(22):4093–4105.

[95] Feinendegen LE. Evidence for beneficial low level radiation effects and radiation hormesis. Br J Radiol. 2005 Jan;78(925):3–7.

[96] Feldkamp LA, Davis LC, Kress JW. Practical cone-beam algorithm. J Opt Soc Am A. 1984;1(6):612–619. Available from: http://josaa.osa.org/abstract.cfm?URI=josaa-1-6-612.

[97] Flynn MJ, Samei E. Experimental comparison of noise and resolution for 2k and 4k storage phosphor radiography systems. Medical Physics. 1999;26(8):1612–1623. Available from: ISI:000082121600022.

[98] Fraass B, Doppke K, Hunt M, Kutcher G, Starkschall G, Stern R, et al. American Association of Physicists in Medicine Radiation Therapy Committee Task Group 53: quality assurance for clinical radiotherapy treatment planning. Med Phys. 1998 Oct;25(10):1773–1829.

[99] Franklin WR. PNPOLY - Point Inclusion in Polygon Test; 2010. Available from: `http://www.ecse.rpi.edu/Homepages/wrf/Research/Short_Notes/pnpoly.html`.

[100] Fujita H, Ueda K, Morishita J, Fujikawa T, Ohtsuka A, Sai T. Basic imaging properties of a computed radiographic system with photostimulable phosphors. Med Phys. 1989;16(1):52–59.

[101] Gao Z, Wilkins D, Eapen L, Morash C, Wassef Y, Gerig L. A study of prostate delineation referenced against a gold standard created from the visible human data. Radiotherapy and Oncology. 2007;85(2):239 – 246. Available from: `http://www.sciencedirect.com/science/article/B6TBY-4PKP47G-1/2/ba5059d333a89631fd2939cb4af2ed24`.

[102] GEMS. CT HiSpeed QX/i Data Sheet rev 1. GE Medical Systems; 2002.

[103] Goitein M, Abrams M. Multi-dimensional treatment planning: I. Delineation of anatomy. Int J Radiat Oncol Biol Phys. 1983 Jun;9(6):777–787.

[104] Goitein M, Abrams M, Rowell D, Pollari H, Wiles J. Multi-dimensional treatment planning: II. Beam's eye-view, back projection, and projection through CT sections. Int J Radiat Oncol Biol Phys. 1983 Jun;9(6):789–797.

[105] Goodenough D. Catphan 500 and 600 Manual. Salem, NY; 2006.

[106] Gray JE, Lisk KG, Haddick DH, Harshbarger JH, Oosterhof A, Schwenker R. Test pattern for video displays and hard-copy cameras. Radiology. 1985 Feb;154(2):519–527.

[107] Greer PB, Popescu CC. Dosimetric properties of an amorphous silicon electronic portal imaging device for verification of dynamic intensity modulated radiation therapy. Med Phys. 2003 Jul;30(7):1618–1627.

[108] Grein EE, Lee R, Luchka K. An investigation of a new amorphous silicon electronic portal imaging device for transit dosimetry. Med Phys. 2002 Oct;29(10):2262–2268.

[109] Grégoire V, Bol A, Geets X, Lee J. Is PET-based treatment planning the new standard in modern radiotherapy? The head and neck paradigm. Semin Radiat Oncol. 2006 Oct;16(4):232–238. Available from: `http://dx.doi.org/10.1016/j.semradonc.2006.04.006`.

[110] Grégoire V, Haustermans K, Geets X, Roels S, Lonneux M. PET-based treatment planning in radiotherapy: a new standard? J Nucl Med. 2007 Jan;48 Suppl 1:68S–77S.

[111] Hanson WF, Berkley LW, Peterson M. Off-axis beam quality change in linear accelerator x-ray beams. Medical Physics. 1980;7(2):145–146. Available from: http://link.aip.org/link/?MPH/7/145/1.

[112] Harrison RM. Doses to organs and tissues from concomitant imaging in radiotherapy: a suggested framework for clinical justification. Br J Radiol. 2008 Dec;81(972):970–974. Available from: http://dx.doi.org/10.1259/bjr/77088613.

[113] Harrison RM. Second cancers following radiotherapy: a suggested common dosimetry framework for therapeutic and concomitant exposures. Br J Radiol. 2004 Dec;77(924):986–990.

[114] Harrison RM, Wilkinson M, Rawlings DJ, Moore M. Doses to critical organs following radiotherapy and concomitant imaging of the larynx and breast. Br J Radiol. 2007 Dec;80(960):989–995. Available from: http://dx.doi.org/10.1259/bjr/32814323.

[115] Harrison RM, Wilkinson M, Shemilt A, Rawlings DJ, Moore M, Lecomber AR. Organ doses from prostate radiotherapy and associated concomitant exposures. Br J Radiol. 2006 Jun;79(942):487–496. Available from: http://dx.doi.org/10.1259/bjr/16187818.

[116] Hassan WMSW. Measurement of Wiener Spectrum of Radiographic Screen-Film Systems. Jurnal Teknologi. 2001 June;34(C):35–42.

[117] Hatcher A. Algebraic Topology. Cambridge: Cambridge University Press; 2002. Available from: http://www.math.cornell.edu/~hatcher/AT/AT.pdf.

[118] Hay GA, Chesters MS. A model of visual threshold detection. Journal of Theoretical Biology. 1977;67(2):221 – 240. Available from: http://www.sciencedirect.com/science/article/B6WMD-4F1SV94-JV/2/5b87ed9a73406d2d9b876237ad3eccee.

[119] Heeger D. Signal Detection Theory. New York; 1997. Available from: http://www.cns.nyu.edu/~david/handouts/sdt-advanced.pdf.

[120] van Herk M, Remeijer P, Rasch C, Lebesque JV. The probability of correct target dosage: dose-population histograms for deriving treatment margins in radiotherapy. International Journal of Radiation Oncology*Biology*Physics. 2000;47(4):1121 – 1135. Available from: http://www.sciencedirect.com/science/article/B6T7X-40HTYJ7-17/2/639208ac3dda2bdbde38fe419e78a60e.

[121] Herman GT. Image Reconstruction from Projections: Fundamentals of Computed Tomography. Academic Press Inc; 1981.

[122] Herman MG. Clinical use of electronic portal imaging. Semin Radiat Oncol. 2005 Jul;15(3):157–167.

[123] Herman MG, Balter JM, Jaffray DA, McGee KP, Munro P, Shalev S, et al. Clinical use of electronic portal imaging: report of AAPM Radiation Therapy Committee Task Group 58. Med Phys. 2001 May;28(5):712–737.

[124] Hillen W, Schiebel U, Zaengel T. Imaging performance of a digital storage phosphor system. Med Phys. 1987;14(5):744–751.

[125] HMSO. The Ionising Radiation (Medical Exposure) Regulations 2000 (Statutory Instrument 2000 No. 1059); 2000. Available from: http://www.opsi.gov.uk/si/si2000/20001059.htm.

[126] HMSO. The Ionising Radiations Regulations 1999 (Statutory Instrument 1999 No. 3232); 2000. Available from: http://www.opsi.gov.uk/si/si1999/19993232.htm.

[127] Holland R, Veling SHJ, Mravunac M, Hendriks JHCL. Histologic multifocality of tis, T1-2 breast carcinomas implications for clinical trials of breast-conserving surgery. Cancer. 1985;56(5):979–990.

[128] Horan G, Roques TW, Curtin J, Barrett A. "Two are better than one": a pilot study of how radiologist and oncologists can collaborate in target volume definition. Cancer Imaging. 2006;6:16–19. Available from: http://dx.doi.org/10.1102/1470-7330.2006.0003.

[129] Hormann K, Agathos A. The point in polygon problem for arbitrary polygons. Computational Geometry. 2001;20(3):131 – 144. Available from: http://www.sciencedirect.com/science/article/B6TYS-448J7J4-3/2/fa70e83c8f1fa667f4bab812b7382936.

[130] Hotelling H. The Generalization of Student's Ratio. The Annals of Mathematical Statistics. 1931;2(3):360–378. Available from: http://www.jstor.org/stable/2957535.

[131] Hounsfield GN. Computerized transverse axial scanning (tomography): Part 1. Description of system. Br J Radiol. 1973;46(552):1016–1022. Available from: http://bjr.birjournals.org/cgi/content/abstract/46/552/1016.

[132] Hubbell J, Seltzer S. Tables of X-Ray Mass Attenuation Coefficients and Mass Energy-Absorption Coefficients (version 1.4).. Gaithersburg, Maryland, USA; 2004. Available from: http://www.nist.gov/pml/data/xraycoef/index.cfm.

[133] IAEA. Transition from 2-D Radiotherapy to 3-D Conformal and Intensity Modulated Radiotherapy: IAEA-TECDOC-1588. Vienna: IAEA; 2008.

[134] IAEA. Setting up a Radiotherapy Programme: Clinical, Medical Physics, Radiation Protection and Safety Aspects. Vienna: IAEA; 2008.

[135] IAEA. Quality assurance in radiotherapy: IAEA-TECDOC-989. Vienna. IAEA; 1997.

[136] IAEA. Radiation dose in radiotherapy from prescription to delivery: IAEA-TECDOC-896. Vienna: IAEA; 1996.

[137] ICRP. ICRP Publication 60: 1990 Recommendations of the International Commission on Radiological Protection. Annals of the ICRP. 1991;21:1–3.

[138] ICRP. Recommendations of the International Commission on Radiological Protection: ICRP Publication 26. ICRP; 1977.

[139] ICRP. Report of the Task Group on Reference Man. ICRP Publication 23. ICRP; 1975.

[140] ICRU. Prescribing, Recording, and Reporting Photon Beam Therapy. ICRU Report 62 (Supplement to ICRU Report 50). Bethesda: International Commission on Radiation Units and Measurements; 1999.

[141] ICRU. Medical Imaging - The Assessment of Image Quality. ICRU Report 54. Bethesda, Maryland, USA: ICRU; 1996.

[142] ICRU. Prescribing, Recording, and Reporting Photon Beam Therapy. ICRU Report 50. Bethesda: International Commission on Radiation Units and Measurements; 1993.

[143] ICRU. Tissue substitutes in radiation dosimetry and measurement. ICRU Report 44. Bethesda: International Commission on Radiation Units and Measurements; 1989.

[144] ICRU. Use of Computers in External Beam Radiotherapy Procedures with High Energy Photons and Electrons. ICRU Report 42. Bethesda: International Commission on Radiation Units and Measurements; 1987.

[145] ICRU. Determination of Absorbed Dose in a Patient Irradiated by Beams of X or Gamma Rays in Radiotherapy Procedures. ICRU Report 24. Bethesda: International Commission on Radiation Units and Measurements; 1976.

[146] IEC. IEC 62220-1:2003. Medical electrical equipment–Characteristics of digital X-ray imaging devices–Part 1: Determination of detective quantum efficiency; 2003.

[147] IEC. IEC 60601-2-44: CT safety standard (Edition2, amendment 1); 2003.

[148] IEC. IEC 61267:1995. Radiation conditions for use in the determination of characteristics for medical diagnostic X-ray equipment.; 1995.

[149] IGE. HiSpeed FX/i Product Data Sheet. IGE Medical Systems; 1998.

[150] ImPACT. CTDosimetry 1.0.3; 2010. Available from: `http://www.impactscan.org/ctdosimetry.htm`.

[151] ImPACT. ImPACT Information Leaflet No. 1: CT Scanner Acceptance Testing; 2001. Available from: `http://www.impactscan.org/download/acceptancetesting.pdf`.

[152] ImPACT. Four Slice CT Scanner Comparison Report, Version 3.04. Medical Devices Agency (ImPACT); 2001. MDA Evaluation Report MDA 01050.

[153] ImPACT. ImPACT Technology Update No. 1, Multi-slice CT scanners. Medical Devices Agency (ImPACT); 2000. MDA Evaluation Report MDA/00/23. Available from: `http://www.impactscan.org/download/acceptancetesting.pdf`.

[154] IPEM. IPEM Report 81: Physics Aspects of Quality Control in Radiotherapy. Mayles WPM, Lake R, McKenzie A, Macaulay EM, Morgan HM, Jordan TJ, et al., editors. York, UK: IPEM; 1999.

[155] IPEM. IPEM Report 96: Guidance for the Clinical Implementation of Intensity Modulated Radiation Therapy. IPEM; 2008.

[156] IPEM. IPEM Report 32: Part III: 2nd Edition Computed Tomography X-Ray Scanners. IPEM; 2003.

[157] Ishida M, Doi K, Loo LN, Metz CE, Lehr JL. Digital image processing: effect on detectability of simulated low-contrast radiographic patterns. Radiology. 1984 Feb;150(2):569–575.

[158] Islam MK, Purdie TG, Norrlinger BD, Alasti H, Moseley DJ, Sharpe MB, et al. Patient dose from kilovoltage cone beam computed tomography imaging in radiation therapy. Med Phys. 2006 Jun;33(6):1573–1582.

[159] Jacobi W. The concept of the effective dose–a proposal for the combination of organ doses. Radiat Environ Biophys. 1975 Jun;12(2):101–109.

[160] Jaffray DA, Drake DG, Moreau M, Martinez AA, Wong JW. A radiographic and tomographic imaging system integrated into a medical linear accelerator for localization of bone and soft-tissue targets. International Journal of Radiation Oncology*Biology*Physics. 1999;45(3):773 – 789. Available from: http://www.sciencedirect.com/science/article/B6T7X-3XM2T3S-17/2/13db4f269996468e4d0af3662f74b0da.

[161] Jaffray DA, Siewerdsen JH, Wong JW, Martinez AA. Flat-panel cone-beam computed tomography for image-guided radiation therapy. International Journal of Radiation Oncology*Biology*Physics. 2002;53(5):1337 – 1349. Available from: http://www.sciencedirect.com/science/article/B6T7X-469C98P-13/2/9a7fce99ad44c2f78808481226fa31e8.

[162] Jefferies S, Taylor A, Reznek R, Party RPW. Results of a national survey of radiotherapy planning and delivery in the UK in 2007. Clin Oncol (R Coll Radiol). 2009 Apr;21(3):204–217. Available from: http://dx.doi.org/10.1016/j.clon.2008.11.017.

[163] Jennings RJ, Jafroudi H, Gagne RM, Fewell TR, Quinn PW, Artz DES, et al. Storage phosphor-based digital mammography using a low-dose x-ray

system optimized for screen-film mammography. Medical Imaging 1996: Physics of Medical Imaging. 1996;2708(1):220–232.

[164] Jhaveri KS, Saini S, Levine LA, Piazzo DJ, Doncaster RJ, Halpern EF, et al. Effect of multislice CT technology on scanner productivity. AJR Am J Roentgenol. 2001 Oct;177(4):769–772.

[165] Jones B. The case for particle therapy. Br J Radiol. 2006 Jan;79(937):24–31. Available from: http://dx.doi.org/10.1259/bjr/81790390.

[166] Jones B, Burnet N. Radiotherapy for the future. BMJ. 2005 Apr;330(7498):979–980. Available from: http://dx.doi.org/10.1136/bmj.330.7498.979.

[167] Judy PF. The line spread function and modulation transfer function of a computed tomographic scanner. Med Phys. 1976;3(4):233–236.

[168] Kalender WA. Computed tomography: fundamentals, system technology, image quality, applications. Springer; 2000.

[169] Kalender WA, Buchenau S, Deak P, Kellermeier M, Langner O, van Straten M, et al. Technical approaches to the optimisation of CT. Phys Med. 2008 Jun;24(2):71–79. Available from: http://dx.doi.org/10.1016/j.ejmp.2008.01.012.

[170] Kalender WA, Polacin A, SÃŒss C. A Comparison of Conventional and Spiral CT: An Experimental Study on the Detection of Spherical Lesions. Journal of Computer Assisted Tomography. 1994;18(2):–. Available from: http://journals.lww.com/jcat/Fulltext/1994/03000/A_Comparison_of_Conventional_and_Spiral_CT__An.1.aspx.

[171] Kan MWK, Leung LHT, Wong W, Lam N. Radiation dose from cone beam computed tomography for image-guided radiation therapy. Int J Radiat Oncol Biol Phys. 2008 Jan;70(1):272–279. Available from: http://dx.doi.org/10.1016/j.ijrobp.2007.08.062.

[172] Karlsson M, Karlsson MG, Nyholm T, Amies C, Zackrisson B. Dedicated Magnetic Resonance Imaging in the Radiotherapy Clinic. International Journal of Radiation Oncology*Biology*Physics. 2009;74(2):644 – 651. Available from: http://www.sciencedirect.com/science/article/B6T7X-4W7DD8G-K/2/32fa41d6d22a5b361f75d5d0fd643fa2.

[173] Kasibhalta M, Yoo S, Yin F, Godfrey D, Oldham M, Das SK. Assessment of Setup Accuracy for Prostate Radiotherapy using On-board Imaging and Cone Beam Computed Tomography. International Journal of Radiation Oncology*Biology*Physics. 2005;63(Supplement 1):S535 – S535. Proceedings of the American Society for Therapeutic Radiology and Oncology 47 th Annual Meeting. Available from: http://www.sciencedirect.com/science/article/B6T7X-4H57X9S-141/2/4f93c30f05f68d93555fdfe738fa5fe9.

[174] Kass M, Witkin A, Terzopoulos D. Snakes: Active contour models. International Journal of Computer Vision. 1988;1:321– 331.

[175] Keat N, Lewis M, Hill A, Edyvean S, Carden J. GE LightSpeed / LightSpeed Plus CT Scanner Technical Evaluation. Medical Devices Agency (ImPACT); 2001. MDA Evaluation Report MDA 01159.

[176] Khoo VS. MRI–"magic radiotherapy imaging" for treatment planning? Br J Radiol. 2000 Mar;73(867):229–233.

[177] Khoo VS, Adams EJ, Saran F, Bedford JL, Perks JR, Warrington AP, et al. A Comparison of clinical target volumes determined by CT and MRI for the radiotherapy planning of base of skull meningiomas. Int J Radiat Oncol Biol Phys. 2000 Mar;46(5):1309–1317.

[178] Kilby W, Sage J, Rabett V. Tolerance levels for quality assurance of electron density values generated from CT in radiotherapy treatment planning. Physics in Medicine and Biology. 2002;47(9):1485. Available from: http://stacks.iop.org/0031-9155/47/i=9/a=304.

[179] King L, Castellano E, McDonagh E. Practical measurement of pre-sampled MTF in digital mammography. In: IPEM Meeting on Quantitative Image Analysis; 2010. .

[180] Kirby M, Carpenter D, Lawrence G, Poynter A, Studdart P. IPEM Report 93: Guidance for Commissioning and QA of a Networked Radiotherapy Department. York: IPEM; 2006.

[181] Kirby MC, Glendinning AG. Developments in electronic portal imaging systems. Br J Radiol. 2006 Sep;79 Spec No 1:S50–S65. Available from: http://dx.doi.org/10.1259/bjr/21517185.

[182] Kirkby C, Sloboda R. Consequences of the spectral response of an a-Si EPID and implications for dosimetric calibration. Med Phys. 2005 Aug;32(8):2649–2658.

[183] Kirkby C, Sloboda R. Comprehensive Monte Carlo calculation of the point spread function for a commercial a-Si EPID. Med Phys. 2005 Apr;32(4):1115–1127.

[184] Kirwin S, Langmack K, Nightingale A. Effect of CT reconstruction kernel and post-processing filter on Hounsfield number constancy in radiotherapy treatment planning. Scope. 2005 March;14(1).

[185] Klein EE, Hanley J, Bayouth J, Yin FF, Simon W, Dresser S, et al. Task Group 142 report: quality assurance of medical accelerators. Med Phys. 2009 Sep;36(9):4197–4212.

[186] Knöös T, Nilsson M, Ahlgren L. A method for conversion of Hounsfield number to electron density and prediction of macroscopic pair production cross-sections. Radiother Oncol. 1986 Apr;5(4):337–345.

[187] Koedooder K, Strackee J, Venema HW. A new method for microdensitometer slit length correction of radiographic noise power spectra. Med Phys. 1986;13(4):469–473.

[188] Kolitsi Z, Dahl O, Loon RV, Drouard J, Dijk JV, Ruden BI, et al. Quality assurance in conformal radiotherapy: DYNARAD consensus report on practice guidelines. Radiother Oncol. 1997 Dec;45(3):217–223.

[189] Kotre CJ. An investigation of search pattern extent in the threshold contrast detection task. Br J Radiol. 2006 May;79(941):437–440. Available from: `http://dx.doi.org/10.1259/bjr/13489819`.

[190] Kulama E. Scanning protocols for multislice CT scanners. Br J Radiol. 2004;77 Spec No 1:S2–S9.

[191] Kupelian PA, Ciezki J, Reddy CA, Klein EA, Mahadevan A. Effect of increasing radiation doses on local and distant failures in patients with localized prostate cancer. Int J Radiat Oncol Biol Phys. 2008 May;71(1):16–22. Available from: `http://dx.doi.org/10.1016/j.ijrobp.2007.09.020`.

[192] Kuriyama K, Onishi H, Sano N, Komiyama T, Aikawa Y, Tateda Y, et al. A new irradiation unit constructed of self-moving

gantry-CT and linac. International Journal of Radiation Oncology*Biology*Physics. 2003;55(2):428 – 435. Available from: http://www.sciencedirect.com/science/article/B6T7X-47MJT3X-K/2/ac6775f4534ed0a56374cc218c53cd1c.

[193] Kutcher GJ, Coia L, Gillin M, Hanson WF, Leibel S, Morton RJ, et al. Comprehensive QA for radiation oncology: report of AAPM Radiation Therapy Committee Task Group 40. Med Phys. 1994 Apr;21(4):581–618.

[194] Kyprianou IS, Rudin S, Bednarek DR, Hoffmann KR. Generalizing the MTF and DQE to include x-ray scatter and focal spot unsharpness: application to a new microangiographic system. Med Phys. 2005 Feb;32(2):613–626.

[195] Langen KM, Pouliot J, Anezinos C, Aubin M, Gottschalk AR, Hsu IC, et al. Evaluation of ultrasound-based prostate localization for image-guided radiotherapy. International Journal of Radiation Oncology*Biology*Physics. 2003;57(3):635 – 644. Available from: http://www.sciencedirect.com/science/article/B6T7X-49MWR6H-5/2/23a76efbe0470158b41725c165f627b6.

[196] Lattanzi J, McNeeley S, Donnelly S, Palacio E, Hanlon A, Schultheiss TE, et al. Ultrasound-based stereotactic guidance in prostate cancer–quantification of organ motion and set-up errors in external beam radiation therapy. Comput Aided Surg. 2000;5(4):289–295. Available from: http://dx.doi.org/3.0.CO;2-D.

[197] Lerski RA, de Certaines JD. II. Performance assessment and quality control in MRI by Eurospin test objects and protocols. Magnetic Resonance Imaging. 1993;11(6):817 – 833. Available from: http://www.sciencedirect.com/science/article/B6T9D-4BY4VD5-182/2/59554dd0299b869c46a785041c2a3e29.

[198] Lerski R, de Wilde J, Boyce D, Ridgway J. IPEM Report 80: Quality Control in Magnetic Resonance Imaging. IPEM; 2000.

[199] Leunens G, Verstraete J, Dutreix A, van der Schueren E. Assessment of dose inhomogeneity at target level by in vivo dosimetry: can the recommended 5target volume be fulfilled in daily practice? Radiother Oncol. 1992 Dec;25(4):242–250.

[200] Lewis M. ImPACT Course Lecture 7: Factors Affecting Patient Dose. ImPACT; 2007. Available from: http://www.impactscan.org/slides/course07/lect7/frame.htm.

[201] Lewis M. Dose measurement in cone beam CT - how appropriate is the 10 cm pencil chamber? 8th CT User Group Meeting; 2006. .

[202] Lewis MA. Multislice CT: opportunities and challenges. Br J Radiol. 2001 Sep;74(885):779–781.

[203] Li X, Yang J, Zhu Y. Digitally reconstructed radiograph generation by an adaptive Monte Carlo method. Physics in Medicine and Biology. 2006;51(11):2745–2752. Available from: http://stacks.iop.org/0031-9155/51/2745.

[204] Liebich J, Licher1 J, Scherf C, Moog J, Kara E, Ramm U. Measurements of the Signal-to-Noise Ratio with "EPID QC Phantom" and "epidSoft 2.0" from PTW. In: World Congress on Medical Physics and Biomedical Engineering, September 7 - 12, 2009, Munich, Germany. vol. Volume 25/1 of Volume 25/1. Springer; 2009. p. 507–510.

[205] Lillicrap SC, Owen B, Williams JR, Williams PC. Code of Practice for high-energy photon therapy dosimetry based on the NPL absorbed dose calibration service. Physics in Medicine and Biology. 1990;35(10):1355–1360.

[206] Lim YW, Lee SU. On the color image segmentation algorithm based on the thresholding and the fuzzy c-means techniques. Pattern Recognition. 1990;23(9):935 – 952. Available from: http://www.sciencedirect.com/science/article/B6V14-48MPM3D-10M/2/5c50622167adf6fdaee9c3b49e5e625c.

[207] Ling CC, Zhang P, Archambault Y, Bocanek J, Tang G, LoSasso T. Commissioning and Quality Assurance of RapidArc Radiotherapy Delivery System. International Journal of Radiation Oncology*Biology*Physics. 2008;72(2):575 – 581. Available from: http://www.sciencedirect.com/science/article/B6T7X-4TF9435-G/2/569f37d2d2de843bd5c225af8558750c.

[208] Low DA, Nystrom M, Kalinin E, Parikh P, Dempsey JF, Bradley JD, et al. A method for the reconstruction of four-dimensional synchronized CT scans acquired during free breathing. Med Phys. 2003 Jun;30(6):1254–1263.

[209] Luchka K, Chen D, Shalev S, Gluhchev G, Rajapakshe R. Assessing radiation and light field congruence with a video based electronic portal imaging device. Medical Physics. 1996;23(7):1245–1252. Available from: http://link.aip.org/link/?MPH/23/1245/1.

[210] Létourneau D, Wong JW, Oldham M, Gulam M, Watt L, Jaffray DA, et al. Cone-beam-CT guided radiation therapy: technical implementation. Radiotherapy and Oncology. 2005;75(3):279 – 286. Available from: http://www.sciencedirect.com/science/article/B6TBY-4G4PBJK-2/2/ba69ed4938e653c301afb7711c23c62e.

[211] Mackie TR, Holmes T, Swerdloff S, Reckwerdt P, Deasy JO, Yang J, et al. Tomotherapy: a new concept for the delivery of dynamic conformal radiotherapy. Med Phys. 1993;20(6):1709–1719.

[212] Mah D, Steckner M, Hanlon A, Freedman G, Milestone B, Mitra R, et al. MRI simulation: effect of gradient distortions on three-dimensional prostate cancer plans. International Journal of Radiation Oncology*Biology*Physics. 2002;53(3):757 – 765. Available from: http://www.sciencedirect.com/science/article/B6T7X-460DP84-13/2/9e02808fa0cf826e229e968249b075d6.

[213] Mah D, Steckner M, Palacio E, Mitra R, Richardson T, Hanks GE. Characteristics and quality assurance of a dedicated open 0.23 T MRI for radiation therapy simulation. Medical Physics. 2002;29(11):2541–2547. Available from: http://link.aip.org/link/?MPH/29/2541/1.

[214] Maidment ADA, Albert M. Conditioning data for calculation of the modulation transfer function. Med Phys. 2003 Feb;30(2):248–253.

[215] Mann HB, Whitney DR. On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. The Annals of Mathematical Statistics. 1947;18(1):50–60. Available from: http://dx.doi.org/10.2307/2236101.

[216] Marsaglia G, Tsang WW, Wang J. Evaluating Kolmogorov's Distribution. Journal of Statistical Software. 2003 11;8(18):1–4. Available from: http://www.jstatsoft.org/v08/i18.

[217] Marshall NW. Calculation of Quantitative Image Quality Parameters: Notes Describing the Use of OBJ_IQ_reduced. Barts and the London NHS Trust: NHS Breast Screen Programme; 2009. 0902.

[218] Marshall NW. A comparison between objective and subjective image quality measurements for a full field digital mammography system. Physics in Medicine and Biology. 2006;51(10):2441–2463. Available from: http://stacks.iop.org/0031-9155/51/2441.

[219] Marshall NW. The practical application of signal detection theory to image quality assessment in x-ray image intensifier-TV fluoroscopy. Phys Med Biol. 2001 Jun;46(6):1631–1649.

[220] Martin CJ. Effective dose: how should it be applied to medical exposures? Br J Radiol. 2007 Aug;80(956):639–647. Available from: http://dx.doi.org/10.1259/bjr/25922439.

[221] Martin CJ, Sharp PF, Sutton DG. Measurement of image quality in diagnostic radiology. Appl Radiat Isot. 1999 Jan;50(1):21–38.

[222] Martin CJ, Sutton DG, Sharp PF. Balancing patient dose and image quality. Appl Radiat Isot. 1999 Jan;50(1):1–19.

[223] Matsinos E. Current status of the CBCT project at Varian Medical Systems. vol. 5745. SPIE; 2005. p. 340–351. Available from: http://link.aip.org/link/?PSI/5745/340/1.

[224] Matsinos E, Kaissl W. The dual-dose imaging technique: a way to enhance the dynamic range of X-ray detectors. Baden, Switzerland; 2006. arXiv:physics/0607024v1 [physics.med-ph]. Available from: http://arxiv.org/abs/physics/0607024v1.

[225] Matsinos E, Kaissl W. The dual-gain mode: a way to enhance the dynamic range of X-ray detectors. Baden, Switzerland; 2006. arXiv:physics/0607024v1 [physics.med-ph]. Available from: http://arxiv.org/abs/physics/0607021v1.

[226] Matsinos E, Kaissl W. The geometric calibration of cone-beam imaging and delivery systems in radiation therapy. Baden, Switzerland; 2006. Available from: http://arxiv.org/abs/physics/0607018v1.

[227] Matsumoto K, Okumura M, Asai Y, Kayuko S, Shintani N, Nishi T, et al. Dosimetric Properties of an Amorphous Silicon Electric Portal Imaging Device for Verification of Dynamic IMRT. In: Proceedings of the Annual Meeting of the American Society for Therapeutic Radiation Oncology (ASTRO); 2008. .

[228] McCullough EC, Holmes TW. Acceptance testing computerized radiation therapy treatment planning systems: direct utilization of CT scan data. Med Phys. 1985;12(2):237–242.

[229] McCurdy BM, Luchka K, Pistorius S. Dosimetric investigation and portal dose image prediction using an amorphous silicon electronic portal imaging device. Med Phys. 2001 Jun;28(6):911–924.

[230] McDermott LN, Louwe RJW, Sonke JJ, van Herk MB, Mijnheer BJ. Dose-response and ghosting effects of an amorphous silicon electronic portal imaging device. Med Phys. 2004 Feb;31(2):285–295.

[231] McDermott LN, Wendling M, Sonke JJ, van Herk M, Mijnheer BJ. Replacing Pretreatment Verification With In Vivo EPID Dosimetry for Prostate IMRT. International Journal of Radiation Oncology*Biology*Physics. 2007;67(5):1568 – 1577. Available from: http://www.sciencedirect.com/science/article/B6T7X-4NC51SM-H/2/cb99bb1d41a500487218a92f931ca9c4.

[232] McGarry CK, Grattan MWD, Cosgrove VP. Optimization of image quality and dose for Varian aS500 electronic portal imaging devices (EPIDs). Phys Med Biol. 2007 Dec;52(23):6865–6877. Available from: http://dx.doi.org/10.1088/0031-9155/52/23/006.

[233] Mcgee KP, Das IJ, Sims C. Evaluation of Digitally Reconstructed Radiographs (Drrs) Used for Clinical Radiotherapy - A Phantom Study. Medical Physics. 1995;22(11):1815–1827. Available from: ISI:A1995TF27800015.

[234] McKenzie A. Modulation transfer function, noise power spectrum and detective quantum efficiency: part 1. In: IPEM Meeting on Quantitative Image Analysis; 2010. .

[235] McKenzie A, Bidmead M, Coffey M, Crellin A, Dobbs J, Driver D, et al. Geometric Uncertainties in Radiotherapy: Defining the Planning Target Volume. British Institute of Radiology; 2003.

[236] McKenzie AL, van Herk M, Mijnheer B. The width of margins in radiotherapy treatment plans. Phys Med Biol. 2000 Nov;45(11):3331–3342.

[237] McParland BJ, Boyd MM. X-ray image intensifier performance and patient doses for combinations of supplemental beam filters and vascular contrast

agents. Physics in Medicine and Biology. 2001;46(1):227–244. Available from: http://stacks.iop.org/0031-9155/46/227.

[238] McRobbie DW, Moore EA, Graves MJ, Prince MR. MRI From Picture to Proton, 2nd Edition. 2nd ed. Cambridge: Cambridge University Press; 2007.

[239] Menon GV, Sloboda RS. Quality assurance measurements of a-Si EPID performance. Med Dosim. 2004;29(1):11–17. Available from: http://dx.doi.org/10.1016/j.meddos.2003.09.002.

[240] MHRA. Display device image quality; 2006. Available from: http://www.mhra.gov.uk/PrintPreview/DefaultSplashPP/CON2018136?DynamicListQuery=&DynamicListSortBy=xCreationDate&DynamicListSortOrder=Desc&DynamicListTitle=&PageNumber=1&Title=Display%20device%20image%20quality&ResultCount=10.

[241] Mijnheer BJ, Battermann JJ, Wambersie A. What degree of accuracy is required and can be achieved in photon and neutron therapy? Radiother Oncol. 1987 Mar;8(3):237–252.

[242] Milan J, Bentley RE. The storage and manipulation of radiation dose data in a small digital computer. Br J Radiol. 1974 Feb;47(554):115–121.

[243] Milickovic N, Baltas D, Giannouli S, Lahanas M, Zamboglou N. CT imaging based digitally reconstructed radiographs and their application in brachytherapy; 2000. Available from: http://stacks.iop.org/0031-9155/45/2787.

[244] Mizowaki T, Cohen GN, Fung AYC, Zaider M. Towards integrating functional imaging in the treatment of prostate cancer with radiation: the registration of the MR spectroscopy imaging to ultrasound/CT images and its implementation in treatment planning. International Journal of Radiation Oncology*Biology*Physics. 2002;54(5):1558 – 1564. Available from: http://www.sciencedirect.com/science/article/B6T7X-4790CSV-1C/2/95476579738536454dca9c5782f3bede.

[245] Mohr GA, Willems P. Factors affecting Probability of Detection with Computed Radiography. In: 17th World Conference on Nondestructive Testing; 2008. .

[246] Moran PR. A physical statistics theory for detectability of target signals in noisy images. I. Mathematical background, empirical review, and development of theory. Med Phys. 1982;9(3):401–413.

[247] Morrison R. The results of treatment of cancer of the bladder-a clinical contribution to radiobiology. Clin Radiol. 1975 Jan;26(1):67–75.

[248] Munro P, Bouius DC. X-ray quantum limited portal imaging using amorphous silicon flat-panel arrays. Med Phys. 1998 May;25(5):689–702.

[249] Munro P, Rawlinson JA, Fenster A. Therapy imaging: a signal-to-noise analysis of a fluoroscopic imaging system for radiotherapy localization. Med Phys. 1990;17(5):763–772.

[250] Murphy MJ, Balter J, Balter S, BenComo JA, Das IJ, Jiang SB, et al. The management of imaging dose during image-guided radiotherapy: report of the AAPM Task Group 75. Med Phys. 2007 Oct;34(10):4041–4063.

[251] Mutch S, Hughes M, Cocker M. Personal Communication: Effective dose from radiotherapy cone-beam CT; 2010.

[252] Mutic S. Quality assurance for computed-tomography simulators and the computed-tomography-simulation process: Report of the AAPM Radiation Therapy Committee Task Group No. 66. Medical Physics. 2003;30:2762–2792.

[253] Mutic S, Palta JR, Butker EK, Das IJ, Huq MS, Loo LND, et al. Quality assurance for computed-tomography simulators and the computed-tomography-simulation process: report of the AAPM Radiation Therapy Committee Task Group No. 66. Med Phys. 2003 Oct;30(10):2762–2792.

[254] Nailon WH, Redpath AT, McLaren DB. Characterisation of radiotherapy planning volumes using textural analysis. Acta Oncologica. 2008;47(7):1303–1308. Available from: http://informahealthcare.com/doi/abs/10.1080/02841860802256467.

[255] NCRP. NCRP Report No. 136 - Evaluation of the Linear-Nonthreshold Dose-Response Model for Ionizing Radiation. NCRP; 2001.

[256] Neal AJ, Sivewright G, Bentley R. Evaluation of a region growing algorithm for segmenting pelvic computed tomography images during radiotherapy planning. Br J Radiol. 1994;67(796):392–395. Available from: http://bjr.birjournals.org/cgi/content/abstract/67/796/392.

[257] Neitzel U, Buhr E, Hilgers G, Granfors PR. Determination of the modulation transfer function using the edge method: influence of scattered radiation. Med Phys. 2004 Dec;31(12):3485–3491.

[258] Neitzel U, Günther-Kohfahl S, Borasi G, Samei E. Determination of the detective quantum efficiency of a digital x-ray detector: comparison of three evaluations using a common image data set. Med Phys. 2004 Aug;31(8):2205–2211.

[259] NHSS. Cancer in Scotland; 2009. Available from: http://www.isdscotland.org/isd/servlet/FileBuffer?namedFile=Cancer_in_Scotland_summary_m.pdf&pContentDispositionType=inline.

[260] Nicolini G, Fogliata A, Vanetti E, Clivio A, Vetterli D, Cozzi L. Testing the GlaaS algorithm for dose measurements on low- and high-energy photon beams using an amorphous silicon portal imager. Med Phys. 2008 Feb;35(2):464–472.

[261] NRAG D. Radiotherapy: Developing a world class service for England. London; 2007. Available from: http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_074575.

[262] Olsen DR, Thwaites DI. Now you see it... Imaging in radiotherapy treatment planning and delivery. Radiotherapy and Oncology. 2007;85(2):173 – 175. Available from: http://www.sciencedirect.com/science/article/B6TBY-4R5H1HJ-5/2/4305ed51701697f8b2a36e5d132a1f95.

[263] Orton EJ, Robar JL. Megavoltage image contrast with low-atomic number target materials and amorphous silicon electronic portal imagers. Phys Med Biol. 2009 Mar;54(5):1275–1289. Available from: http://dx.doi.org/10.1088/0031-9155/54/5/012.

[264] Padgett R, Kotre CJ. Development and application of programs to measure modulation transfer function, noise power spectrum and detective quantum efficiency. Radiat Prot Dosimetry. 2005;117(1-3):283–287. Available from: http://dx.doi.org/10.1093/rpd/nci740.

[265] Parker RP, Hobday PA, Cassell KJ. The direct use of CT numbers in radiotherapy dosage calculations for inhomogeneous media. Phys Med Biol. 1979 Jul;24(4):802–809.

[266] Pascoal A, Lawinski CP, Honey I, Blake P. Evaluation of a software package for automated quality assessment of contrast detail images—comparison with subjective visual assessment. Physics in Medicine and Biology. 2005;50(23):5743–5757. Available from: http://stacks.iop.org/0031-9155/50/5743.

[267] Pawlowski JM, Yang ES, Malcolm AW, Coffey CW, Ding GX. Reduction of dose delivered to organs at risk in prostate cancer patients via image-guided radiation therapy. Int J Radiat Oncol Biol Phys. 2010 Mar;76(3):924–934. Available from: http://dx.doi.org/10.1016/j.ijrobp.2009.06.068.

[268] Perez CA, Stanley K, Rubin P, Kramer S, Brady L, Perez-Tamayo R, et al. A prospective randomized study of various irradiation doses and fractionation schedules in the treatment of inoperable non-oat-cell carcinoma of the lung. Preliminary report by the Radiation Therapy Oncology Group. Cancer. 1980 Jun;45(11):2744–2753.

[269] Pesznyák C, Fekete G, Mózes A, Kiss B, Király R, Polgár I, et al. Quality Control of Portal Imaging with PTW EPID QC PHANTOM. Strahlenther Onkol. 2009 Jan;185(1):56–60. Available from: http://dx.doi.org/10.1007/s00066-009-1905-z.

[270] Pieterman RM, van Putten JW, Meuzelaar JJ, Mooyaart EL, Vaalburg W, Koëter GH, et al. Preoperative staging of non-small-cell lung cancer with positron-emission tomography. N Engl J Med. 2000 Jul;343(4):254–261.

[271] Platten D. The effect of experimental setup on the measurement of MTF and NNPS. In: IPEM Meeting on Quantitative Image Analysis; 2010. .

[272] Poletti JL, McLean D. The effect of source to image-receptor distance on effective dose for some common X-ray projections. Br J Radiol. 2005 Sep;78(933):810–815. Available from: http://dx.doi.org/10.1259/bjr/74823655.

[273] Pouliot J, Bani-Hashemi A, Chen J, Svatos M, Ghelmansarai F, Mitschke M, et al. Low-dose megavoltage cone-beam CT for radiation therapy. International Journal of Radiation Oncology*Biology*Physics. 2005;61(2):552 – 560. Available from: http://www.sciencedirect.com/science/article/B6T7X-4F9MJMX-14/2/dd1aa02ad745de9702c0760e5ed488ea.

[274] Press WH. Numerical recipes : the art of scientific computing. 3rd ed. Cambridge University Press; 2007.

[275] PTW. User Manual: EPID QC Phantom and epidSoft Software (Software Version 2.0 or Higher); 2006.

[276] Purchasing N, Agency S. Protocol: Tomographic Image Guided Radiotherapy CEP061212-135 Draft for Stakeholder Comment. NHS Purchasing and Supply Agency - Centre for Evidence-based Purchasing; 2008.

[277] Purdy JA. Future directions in 3-d treatment planning and delivery: a physicist's perspective. International Journal of Radiation Oncology*Biology*Physics. 2000;46(1):3 – 6. Available from: http://www.sciencedirect.com/science/article/B6T7X-3Y8VJ67-2/2/45b1049a2bd7b7bd7753b3aa996fc5c5.

[278] Purwar AK. Preliminary Evaluation of the QCkV-1 Phantom and CyberKnife Digital Flat Panel Detector Imager Response. In: Proceedings of the Annual Meeting of the AAPM; 2008. .

[279] Raaijmakers AJE, Raaymakers BW, van der Meer S, Lagendijk JJW. Integrating a MRI scanner with a 6 MV radiotherapy accelerator: impact of the surface orientation on the entrance and exit dose due to the transverse magnetic field. Phys Med Biol. 2007 Feb;52(4):929–939. Available from: http://dx.doi.org/10.1088/0031-9155/52/4/005.

[280] Raaymakers BW, Raaijmakers AJE, Kotte ANTJ, Jette D, Lagendijk JJW. Integrating a MRI scanner with a 6 MV radiotherapy accelerator: dose deposition in a transverse magnetic field. Physics in Medicine and Biology. 2004;49(17):4109–4118. Available from: http://stacks.iop.org/0031-9155/49/4109.

[281] Rajapakshe R, Luchka K, Shalev S. A quality control test for electronic portal imaging devices. Medical Physics. 1996;23(7):1237–1244. Available from: http://link.aip.org/link/?MPH/23/1237/1.

[282] Ranger NT, Samei E, Dobbins JT, Ravin CE. Assessment of detective quantum efficiency: intercomparison of a recently introduced international standard with prior methods. Radiology. 2007 Jun;243(3):785–795. Available from: http://dx.doi.org/10.1148/radiol.2433060485.

[283] RCR. A Guide to Understanding the Implications of the Ionising Radiation (Medical Exposure) Regulations in Radiotherapy. London: Royal College of Radiologists; 2008.

[284] RCR. On target: ensuring geometric accuracy in radiotherapy. London: Royal College of Radiologists; 2008.

[285] RCR. Towards Safer Radiotherapy. London: Royal College of Radiologists; 2008.

[286] RCR. Radiotherapy Dose-Fractionation. London: Royal College of Radiologists; 2006.

[287] RCR. Imaging for Oncology: Collaboration between Clinical Radiologists and Clinical Oncologists in Diagnosis, Staging and Radiotherapy Planning. London: Royal College of Radiologists; 2004.

[288] RCR. Development and Implementation of Conformal Radiotherapy in the United Kingdom; 2002. Available from: http://www.rcr.ac.uk/docs/oncology/pdf/Conformal_Radiotherapy.pdf.

[289] Reilly AJ, Erridge SC, Ironside J, Little F, Junor E, McLaren D, et al. A Novel Tool for Assessing Set-up Uncertainties Applied to Radical Radiotherapy Patients. Clinical Oncology. 2007;19(3):S6 – S7. Available from: http://www.sciencedirect.com/science/article/B6WXW-4N3GNSC-C/2/8e0bd918e4d1692a1b9322762bd41c7a.

[290] Reilly AJ, MacLeod AS, Thwaites DI. Digitally Reconstructed Radiographs: A New Phantom for Verifying Accuracy and Optimising Quality. Clinical Oncology. 2007;19(3):S43 – S44. Available from: http://www.sciencedirect.com/science/article/B6WXW-4N3GNSC-4V/2/20f7255bddf7d1e18d652f332de582a8.

[291] Reilly A, Sutton D. IPEM Report 78: Spectrum Processor Software; 1997.

[292] Reilly AJ, Thwaites DI. Automated Analysis Software for the Objective Assessment and Optimisation of Radiotherapy Image Quality. International Journal of Radiation Oncology*Biology*Physics. 2007;69(3, Supplement 1):S719 – S719. Proceedings of the American Society for Therapeutic Radiology and Oncology 49th Annual Meeting, 49th Annual Meeting of the American Society for Therapeutic Radiology and

Oncology. Available from: http://www.sciencedirect.com/science/article/B6T7X-4PS4X0B-1M7/2/b190d0e5933f5eba89c1d39d919a84d8.

[293] Reilly AJ. Evaluation of AutoQA Lite Image Quality Measurement Software. In: 5th Meeting of the CT User Group; 2004. .

[294] Reilly AJ. Grade A Training Portfolio: Magnetic Resonance Imaging. Grampian University Hospitals NHS Trust; 2001.

[295] Reilly AJ, Skrzynski W, Thwaites DI. EPID Quality Assurance: A Fresh Approach. Clinical Oncology. 2005;17(Supplement 1):S19.

[296] Reilly AJ, Thwaites DI. IQ Works: An Automated Image Analysis Framework for the Objective Assessment and Optimisation of Image Quality. RSNA; 2007. .

[297] Reilly AJ, Weir N. A New Package and the Evaluation of Detectors for Display and Projector Luminance Response Curve Calibration against the DICOM Part 14 Grayscale Standard Display Function (GSDF). RSNA; 2007. .

[298] Reilly AJ, Weir N. Automated analysis of Catphan 500 and 600 images from single and multi-slice CT scanners. In: 7th Meeting of the CT User Group; 2006. .

[299] Robertson JM, Haken RKT, Hazuka MB, Turrisi AT, Martel MK, Pu AT, et al. Dose escalation for non-small cell lung cancer using conformal radiation therapy. Int J Radiat Oncol Biol Phys. 1997 Mar;37(5):1079–1085.

[300] Robinson P, Collins CD, Ryder WD, Carrington BM, Hutchinson CE, Bell D, et al. Relationship of MRI and clinical staging to outcome in invasive bladder cancer treated by radiotherapy. Clin Radiol. 2000 Apr;55(4):301–306. Available from: http://dx.doi.org/10.1053/crad.1999.0381.

[301] ROSE A. The sensitivity performance of the human eye on an absolute scale. J Opt Soc Am. 1948 Feb;38(2):196–208.

[302] Rose A. A Unified Approach to the Performance of Photographic Film, Television Pickup Tubes, and the Human Eye. Journal of the Society of Motion Picture Engineers. 1946;47:273–294.

[303] Rose A. The Relative Sensitivities of Television Pickup Tubes, Photographic Film, and the Human Eye. Proceedings of the IRE. 1942 June;30(6):293–300.

[304] Rosenman JG, Miller EP, Tracton G, Cullip TJ. Image registration: an essential part of radiation therapy treatment planning. International Journal of Radiation Oncology*Biology*Physics. 1998;40(1):197 – 205. Available from: http://www.sciencedirect.com/science/article/B6T7X-3S3MRF2-11/2/c630e679edab780f98736162d01c3765.

[305] Samei E, Badano A, Chakraborty D, Compton K, Cornelius C, Corrigan K, et al. Assessment of display performance for medical imaging systems: executive summary of AAPM TG18 report. Med Phys. 2005 Apr;32(4):1205–1225.

[306] Samei E, Buhr E, Granfors P, Vandenbroucke D, Wang XH. Comparison of edge analysis techniques for the determination of the MTF of digital radiographic systems. Physics in Medicine and Biology. 2005;50(15):3613–3625. Available from: ISI:000231321600009.

[307] Samei E, Flynn MJ, Reimann DA. A method for measuring the presampled MTF of digital radiographic systems using an edge test device. Med Phys. 1998 Jan;25(1):102–113.

[308] Samei E, Ranger NT, Dobbins JT, Chen Y. Intercomparison of methods for image quality characterization. I. Modulation transfer function. Medical Physics. 2006 May;33(5):1454–1465.

[309] Sandrik JM, Wagner RF. Absolute measures of physical image quality: Measurement and application to radiographic magnification. Medical Physics. 1982;9(4):540–549. Available from: http://link.aip.org/link/?MPH/9/540/1.

[310] Sandrik JM, Wagner RF. Radiographic screen-film noise power spectrum: variation with microdensitometer slit length. Appl Opt. 1981;20(16):2795–2798.

[311] Sawant A, Antonuk L, El-Mohri Y. Slit design for efficient and accurate MTF measurement at megavoltage x-ray energies. Med Phys. 2007 May;34(5):1535–1545.

[312] Sawyer L. Calculation of effective doses for radiotherapy cone-beam CT and nuclear medicine Hawkeye CT. In: CT User Group Meeting; 2006. .

[313] SAWYER LJ, WHITTLE SA, MATTHEWS ES, STARRITT HC, JUPP TP. Estimation of organ and effective doses resulting from cone beam CT imaging for radiotherapy treatment planning. Br J Radiol. 2009;82(979):577–584. Available from: http://bjr.birjournals.org/cgi/content/abstract/82/979/577.

[314] Schneider U, Pedroni E, Lomax A. The calibration of CT Hounsfield units for radiotherapy treatment planning. Phys Med Biol. 1996 Jan;41(1):111–124.

[315] Schnitzler AD. Image-detector model and parameters of the human visual system. J Opt Soc Am. 1973 Nov;63(11):1357–1368.

[316] SEHD. Cancer in Scotland: Radiotherapy Activity Planning for Scotland 2011-2015; 2006. Available from: http://www.scotland.gov.uk/Resource/Doc/90297/0021749.pdf.

[317] Sharp PF. Quantifying image quality. Clin Phys Physiol Meas. 1990;11 Suppl A:21–26.

[318] Sharp PF. The Problem of Measuring Image Quality in the Clinical Environment. The Journal of Photographic Science. 1989;37:60–64.

[319] Shih HA, Jiang SB, Aljarrah KM, Doppke KP, Choi NC. Internal target volume determined with expansion margins beyond composite gross tumor volume in three-dimensional conformal radiotherapy for lung cancer. Int J Radiat Oncol Biol Phys. 2004 Oct;60(2):613–622. Available from: http://dx.doi.org/10.1016/j.ijrobp.2004.05.031.

[320] Shirato H, Shimizu S, Kunieda T, Kitamura K, van Herk M, Kagei K, et al. Physical aspects of a real-time tumor-tracking system for gated radiotherapy. International Journal of Radiation Oncology*Biology*Physics. 2000;48(4):1187 – 1195. Available from: http://www.sciencedirect.com/science/article/B6T7X-41KP2BW-16/2/5c9fb01765c90fa1611af5297fa2dc4c.

[321] Shukovsky LJ. Dose, time, volume relationships in squamous cell carcinoma of the supraglottic larynx. Am J Roentgenol Radium Ther Nucl Med. 1970 Jan;108(1):27–29.

[322] Siebers JV, Kim JO, Ko L, Keall PJ, Mohan R. Monte Carlo computation of dosimetric amorphous silicon electronic portal images. Med Phys. 2004 Jul;31(7):2135–2146.

[323] Sims R, Isambert A, GrÃ©goire V, Bidault F, Fresco L, Sage J, et al.. A pre-clinical assessment of an atlas-based automatic segmentation tool for the head and neck. Elsevier Scientific Publishers; 2009. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0167814009004496.

[324] Skrzynski W, Reilly AJ, Thwaites DI, Bulski W. EPID Image Quality: Las Vegas Phantom as an Objective Tool. 13th Congress of the Polish Society of Medical Physics 2005; 2005. .

[325] Smith AR. Proton therapy. Physics in Medicine and Biology. 2006;51(13):R491–R504. Available from: http://stacks.iop.org/0031-9155/51/R491.

[326] Soete G, de Steene JV, Verellen D, Vinh-Hung V, den Berge DV, Michielsen D, et al. Initial clinical experience with infrared-reflecting skin markers in the positioning of patients treated by conformal radiotherapy for prostate cancer. International Journal of Radiation Oncology*Biology*Physics. 2002;52(3):694 – 698. Available from: http://www.sciencedirect.com/science/article/B6T7X-453H73M-J/2/3f54a67377bfcf4747cbf3d236926e56.

[327] Sokole EB. IAEA quality control atlas for scintillation camera systems. Vienna. IAEA; 2003.

[328] Song WY, Kamath S, Ozawa S, Ani SA, Chvetsov A, Bhandare N, et al. A dose comparison study between XVI and OBI CBCT systems. Med Phys. 2008 Feb;35(2):480–486.

[329] Stenman A, Gustafsson F. Adaptive smoothing methods for frequency-function estimation. Automatica. 2001;37(5):675 – 685.

[330] Stratford J, Ball K, Henry AM, Cullen JN, Swindell R, Price P, et al. Radiotherapy treatment verification in the UK: an audit of practice in 2004. Clin Oncol (R Coll Radiol). 2006 Feb;18(1):15–22.

[331] Swamy K, Narayanan VKS, Basu S, Maiya V, Achari R, Deshpande S, et al. Dose escalation in image-guided, intensity-modulated radiotherapy of carcinoma prostate: initial experience in India. J Cancer Res

Ther. 2009;5(4):277–283. Available from: http://dx.doi.org/10.4103/0973-1482.59909.

[332] Tapiovaara M. Relationships between physical measurements and user evaluation of iamge quality in medical radiology - A review. Finish Radiation and Nuclear Safety Authority; 2006. STUK-A219. Available from: http://www.stuk.fi/julkaisut/stuk-a/stuk-a219.pdf.

[333] Tapiovaara M. Objective measurement of image quality in fluoroscopic x-ray equipment: FluoroQuality. Finish Radiation and Nuclear Safety Authority; 2003. STUK-A196. Available from: http://www.stuk.fi/julkaisut/stuk-a/stuk-a196.pdf.

[334] Tapiovaara MJ. Review of relationships between physical measurements and user evaluation of image quality. Radiat Prot Dosimetry. 2008;129(1-3):244–248. Available from: http://dx.doi.org/10.1093/rpd/ncn009.

[335] Tapiovaara MJ, Wagner RF. SNR and noise measurements for medical imaging: I. A practical approach based on statistical decision theory. Phys Med Biol. 1993 Jan;38(1):71–92.

[336] Thomas SJ. Relative electron density calibration of CT scanners for radiotherapy treatment planning. Br J Radiol. 1999 Aug;72(860):781–786.

[337] Thorson T, Prosser T. X-ray volume imaging in image-guided radiotherapy. Medical Dosimetry. 2006;31(2):126 – 133. Image-Guided Radiation Therapy, Part 2: Cone-Beam Imaging and Respiratory Motion. Available from: http://www.sciencedirect.com/science/article/B6T9J-4JXB0VW-7/2/83d5f7840da38546af5bc54bba58e8c1.

[338] Umbaugh SE. Computer imaging: digital image analysis and processing. CRC Press; 2005. Available from: http://www.ee.siue.edu/CVIPtools/.

[339] Vannier MW. 18: Tumor Imaging, Analysis and Treatment Planning. In: Jacob Beutel RLV Harold L Kundel, editor. Handbook of Medical Imaging. 1st ed. SPIE Publications; 2000. p. 1131–1151.

[340] Verduijn GM, Bartels LW, Raaijmakers CPJ, Terhaard CHJ, Pameijer FA, van den Berg CAT. Magnetic Resonance Imaging Protocol Optimization for Delineation of Gross Tumor Volume in Hypopharyngeal and Laryngeal Tumors. International Journal of Radiation Oncology*Biology*Physics. 2009;74(2):630 – 636. Available from:

http://www.sciencedirect.com/science/article/B6T7X-4W7DD8G-H/2/1736224f4526dc357c42add332178ea8.

[341] Verellen D, Ridder MD, Linthout N, Tournel K, Soete G, Storme G. Innovations in image-guided radiotherapy. Nat Rev Cancer. 2007 Dec;7(12):949–960. Available from: http://dx.doi.org/10.1038/nrc2288.

[342] Verellen D, Soete G, Linthout N, Acker SV, Roover PD, Vinh-Hung V, et al. Quality assurance of a system for improved target localization and patient set-up that combines real-time infrared tracking and stereoscopic X-ray imaging. Radiotherapy and Oncology. 2003;67(1):129 – 141. Available from: http://www.sciencedirect.com/science/article/B6TBY-4834NC0-1/2/a912a0f14c14cefae401d4c872961a69.

[343] Vetterli D, Riem H, Aebersold DM, Greiner RH, Manser P, Cossmann P, et al. Introduction of a novel dose saving acquisition mode for the PortalVision aS500 EPID to facilitate on-line patient setup verification. Med Phys. 2004 Apr;31(4):828–831.

[344] VMS. Specifications: On-Board Imager (RAD 9502G). Palo Alto, CA, USA; 2008.

[345] VMS. On-board Imager Reference Guide (OBI) Advanced Imaging. Palo Alto, USA; 2008.

[346] VMS. Acuity ConeBeam CT Reference Guide. Palo Alto, California, USA; 2007.

[347] VMS. Image Acquisition System 3 Reference Guide. Palo Alto, USA; 2007.

[348] VMS. Vision XimaVision / ScanVision 6.5 Reference Guide. Palo Alto, USA; 2003.

[349] VMS. Palo Alto;.

[350] VMS. PortalVision aS500 Rel 6 Reference Manual. Palo Alto, USA; 2000.

[351] Waddington SP, McKenzie AL. Assessment of effective dose from concomitant exposures required in verification of the target volume in radiotherapy. Br J Radiol. 2004 Jul;77(919):557–561.

[352] Wilcoxon F. Individual Comparisons by Ranking Methods. Biometrics Bulletin. 1945;1(6):80–83. Available from: http://dx.doi.org/10.2307/3001968.

[353] Williams MB, Mangiafico PA, Simoni PU. Noise power spectra of images from digital mammography detectors. Med Phys. 1999 Jul;26(7):1279–1293.

[354] Williams MB, Yaffe MJ, Maidment ADA, Martin MC, Seibert JA, Pisano ED. Image quality in digital mammography: image acquisition. J Am Coll Radiol. 2006 Aug;3(8):589–608. Available from: http://dx.doi.org/10.1016/j.jacr.2006.04.004.

[355] Workman A, Cowen AR. Signal, noise and SNR transfer properties of computed radiography. Physics in Medicine and Biology. 1993;38(12):1789–1808. Available from: http://stacks.iop.org/0031-9155/38/1789.

[356] Wowk B, Radcliffe T, Leszczynski KW, Shalev S, Rajapakshe R. Optimization of metal/phosphor screens for on-line portal imaging. Med Phys. 1994 Feb;21(2):227–235.

[357] Wu S, Jin X, Xie C, Cao G. Optimal steel thickness combined with computed radiography for portal imaging of nasopharyngeal cancer patients. Med Phys. 2005 Oct;32(10):3112–3116.

[358] Yin FF, Guan H, Lu W. A technique for on-board CT reconstruction using both kilovoltage and megavoltage beam projections for 3D treatment verification. Medical Physics. 2005;32(9):2819–2826. Available from: http://link.aip.org/link/?MPH/32/2819/1.

[359] Yoo S, Kim GY, Hammoud R, Elder E, Pawlicki T, Guan H, et al. A quality assurance program for the on-board imager. Medical Physics. 2006;33(11):4431–4447. Available from: http://link.aip.org/link/?MPH/33/4431/1.

[360] Yoo T, Ackerman M, Lorensen W, Schroeder W, Chalana V, Aylward S, et al. Engineering and Algorithm Design for an Image Processing API: A Technical Report on ITK - the Insight Toolkit. 2002 01;.

[361] Yu MK, Sloboda RS, Murray B. Linear accelerator photon beam quality at off-axis points. Medical Physics. 1997;24(2):233–239. Available from: http://link.aip.org/link/?MPH/24/233/1.

[362] Zhu L, Xie Y, Wang J, Xing L. Scatter correction for cone-beam CT in radiation therapy. Medical Physics. 2009;36(6):2258–2268. Available from: http://link.aip.org/link/?MPH/36/2258/1.

[363] Allisy-Roberts P, editor. Medical and Dental Guidance Notes: A good practice guide on all aspects of ionising radiation protection in the clinical environment. York: IPEM; 2002.

[364] Brahme A, editor. Accuracy requirements and quality assurance of external beam therapy with photons and electrons. vol. Supplement 1. Acta Oncol; 1988.

[365] Williams JR, Thwaites DI, editors. Radiotherapy physics in practice. 2nd ed. Oxford University Press; 2000.

[366] IKVM Website;. Available from: http://www.ikvm.net/.

[367] MeVisLab Website;. Available from: http://www.mevislab.de.

[368] Artinis Website; 2010. Available from: http://www.artinis.com/.

[369] AutoQA Lite Website; 2010. Available from: http://www.theiris.com/autoqalite/index.html.

[370] PTW epidSoft Website; 2010. Available from: http://www.ptw.de/press_epid_qc.html.

[371] IDL. Discover What's In Your Data.; 2010. Available from: http://www.ittvis.com/portals/0/pdfs/idl/idl_brochure.pdf.

[372] PIPSpro Website; 2010. Available from: http://www.standardimaging.com/product_home.php?id=14.

[373] Quick MTF Website; 2010. Available from: http://www.quickmtf.com/.

[374] Digital Imaging and Communications in Medicine (DICOM) Part 14: Grayscale Standard Display Function. Rosslyn, Virginia, USA; 2009. Available from: ftp://medical.nema.org/medical/dicom/2009/09_14pu.pdf.

[375] Eclipse Version 8.9 Customer Release Notes. Palo Alto; 2009.

[376] ICRP 110: Adult Reference Computational Phantoms. Annals of the ICRP. 2009;39(2):IFC – IFC.

[377] AAPM Report No 96: The Measurement, Reporting, and Management of Radiation Dose in CT - Report of AAPM Task Group 23: CT Dosimetry. AAPM; 2008.

[378] The 2007 Recommendations of the International Commission on Radiological Protection. ICRP publication 103. Ann ICRP. 2007;37(2-4):1–332. Available from: `http://dx.doi.org/10.1016/j.icrp.2007.10.003`.

[379] S-Plus 8 Guide to Statistics, Volume 2. Seattle: Insightful Corporation; 2007.

[380] Vision Applications Rel 6.5. System Administrators Manual. Palo Alto; 2003.

[381] PIPSpro User's Guide. vol. Version 3.2.2. The Raincoast Executive Centre, 201 Selby Street, Nanaimo, British Columbia. V9R 2R2. Canada.: Masthead Imaging Corporation; 2002.

[382] SMPTE Recommended Practice: Specifications for Medical Diagnostic Imaging Test Pattern for Television Monitors and Hard-Copy Recording Cameras. Society of Motion Picture and Television Engineers; 1991. RP 133-1991.